



# STIC Search Report

## Biotech-Chem Library

STIC Database Tracking Number: 121495

To: Sarvamangala Devi  
Location: REM 3C18  
Art Unit: 1645  
Thursday, May 13, 2004  
  
Case Serial Number: 10/089787

From: Beverly Shears  
Location: Remsen Bldg.  
RM 1A54  
Phone: 571-272-2528  
  
beverly.shears@uspto.gov

### Shears, Beverly

From: Devi, Sarvamangala  
Sent: Thursday, May 06, 2004 4:27 PM  
To: Shears, Beverly  
Subject: 10/089,787

Beverly:

Please perform a sequence and an interference search for SEQ ID NO: 1 and SEQ ID NO: 2 in application SN 10/089,787.

Thanx.

S. DEVI, Ph.D.  
AU 1645  
Rems - 3C18

### STAFF USE ONLY

Date completed: 05-13-04  
Searcher: Beverly C 2528  
Terminal time: 20  
Elapsed time:  
CPU time:  
Total time: 25  
Number of Searches:  
Number of Databases: 1

Search Site  
☐ STIC  
☐ CM-1  
☐ Pre-S  
  
Type of Search  
☐ N.A. Sequence  
☐ A.A. Sequence  
☐ Structure  
☐ Bibliographic

Vendors  
☐ IG  
☐ STN  
☐ Dialog  
☐ APS  
☐ Geninfo  
☐ SDC  
☐ DARC/Questel  
☒ Other CGN

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2004, 08:37:38 ; Search time 46 Seconds  
(without alignments)  
2469.273 Million cell updates/sec

Title: US-10-089-787-2

Perfect score: 1837

Sequence: 1 MLLLAELYQFYKFGVGFQ.....RFMIITVLVLGLATLKLK 360

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SPTREMBL\_25.\*

- 1: sp\_archae.\*
- 2: sp\_bacteria.\*
- 3: sp\_fungi.\*
- 4: sp\_human.\*
- 5: sp\_invertebrate.\*
- 6: sp\_mammal.\*
- 7: sp\_mhc.\*
- 8: sp\_organelle.\*
- 9: sp\_phage.\*
- 10: sp\_plant.\*
- 11: sp\_rodent.\*
- 12: sp\_virus.\*
- 13: sp\_vertebrate.\*
- 14: sp\_unclassified.\*
- 15: sp\_virus.\*
- 16: sp\_bacteriap.\*
- 17: sp\_archae.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	ID	Description
1	1689	91.9	Q8n79	pseudomonas
2	1658	90.3	Q87W2	pseudomonas
3	1314.5	71.6	Q8ZVS6	Q8ZVS6 nitrosomona
4	1306	71.1	Q8E9P5	Q8E9P5 shewanella
5	1296.5	70.6	Q83F26	Q83F26 coxiella bu
6	1212.5	66.0	Q8XVI4	Q8XVI4 ralstonia s
7	1186	64.6	Q7VP57	Q7VP57 haemophilus
8	1184.5	64.5	Q7WFR9	Q7WFR9 bordetella
9	1184.5	64.5	Q7VUQ0	Q7VUQ0 bordetella
10	1181.5	64.3	Q7W4B1	Q7W4B1 bordetella
11	976	53.1	Q89FU4	Q89FU4 bradyrhizob
12	830.5	45.2	Q7VQJ0	Q7VQJ0 candidatus
13	828	45.1	Q8KGD1	Q8KGD1 chlorobium
14	819	44.6	Q8P4J3	Q8P4J3 leptospira
15	782.5	42.6	Q8F172	Q8F172 campylobact
16	704.5	38.4	Q7VGZ9	Q7VGZ9 helicobacte

17	681.5	37.1	326	16	Q8ER51	Q8ER51 oceanobacil
18	671	36.5	324	16	Q81WC8	Q81WC8 bacillus an
19	661.5	36.0	318	2	Q8GE12	Q8GE12 hellobacill
20	661	36.0	324	16	Q819Q1	Q819Q1 bacillus ce
21	652	35.5	319	16	Q894B9	Q894B9 clostridium
22	639	34.8	321	16	Q8CPK7	Q8CPK7 staphylococ
23	638	34.7	422	16	Q8A255	Q8A255 bacteroides
24	617.5	33.6	357	16	Q82AD9	Q82AD9 streptomyc
25	615.5	33.5	316	16	Q8E9G3	Q8E9G3 thermomace
26	611	33.3	321	16	Q88V79	Q88V79 lactobacill
27	572	31.1	368	16	Q8G4Q7	Q8G4Q7 bifidobacte
28	555	30.2	373	16	Q7V3S7	Q7V3S7 prochloroco
29	533.5	29.0	369	16	Q83GN4	Q83GN4 tropheryma
30	532.5	29.0	369	16	Q83HJ8	Q83HJ8 tropheryma
31	526.5	28.7	339	16	Q8DVM4	Q8DVM4 streptococ
32	524	28.5	359	16	Q7UZF8	Q7UZF8 streptococ
33	522.5	28.4	378	16	Q8DK95	Q8DK95 synecococ
34	521	28.4	368	16	Q7U3B6	Q7U3B6 synecococ
35	517.5	28.2	370	16	Q7V9F5	Q7V9F5 prochloroco
36	508.5	27.7	336	16	Q8E779	Q8E779 streptococ
37	508.5	27.7	336	16	Q8E1R5	Q8E1R5 streptococ
38	449.5	24.5	312	16	Q9RTD0	Q9RTD0 deinococcus
39	400	21.8	326	10	Q8L7I8	Q8L7I8 arabidopsis
40	284	15.5	356	17	Q8TXC1	Q8TXC1 methanopyru
41	254.5	13.9	313	10	Q9MOM0	Q9MOM0 arabidopsis
42	247	13.4	298	17	Q8U3R2	Q8U3R2 pyrococcus
43	246.5	13.4	351	16	Q92D62	Q92D62 listeria in
44	245	13.3	363	16	Q8CX79	Q8CX79 oceanobacil
45	240.5	13.1	357	16	Q81X14	Q81X14 bacillus an

#### ALIGNMENTS

#### RESULT 1

Q88N79 ID Q88N79 PRELIMINARY; PRT; 360 AA.  
AC Q88N79; 01-JUN-2003 (TREMBLrel. 24, Created)  
DT 01-JUN-2003 (TREMBLrel. 24, Last sequence update)  
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)  
DE Phospho-N-acetylmuramoyl-pentapeptide- transferase.  
GN MRAY OR Pp1334.  
OS Pseudomonas putida (strain KT2440).  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;  
OC Pseudomonadaceae; Pseudomonas.  
OX NCBITaxID=160488;  
[1]  
RN SEQUENCE FROM N.A.  
RX MEDLINE=22423060; PubMed=12534463;  
RA Nelson K.E., Weinel C., Paulsen I.T., Dodson R.J., Hilbert H.,  
RA Martins dos Santos V.A.P., Fouts D.E., Gill S.R., Pop M., Holmes M.,  
RA Brinkac L., Beanan M., DeBoy R.T., Daugherty S., Kolonay J.,  
RA Madupu R., Nelson W., White O., Peterson J., Khouri H., Hance I.,  
RA Chris Lee P., Holtzapfle E., Scanlan D., Tran K., Moazzez A.,  
RA Utterback T., Rizzo M., Lee K., Kosack D., Moestl D., Wedler H.,  
RA Lauber J., Sjepeandic D., Hoheisel J., Straetz M., Heim S.,  
RA Kiewitz C., Eisen J., Timmis K.N., Duesterhoeft A., Tuemmler B.,  
RA Fraser C.M.;  
RA "Complete genome sequence and comparative analysis of the  
RT metabolically versatile Pseudomonas putida KT2440.";  
RL Environ. Microbiol. 4:799-808 (2002).  
DR EMBL; AE016779; AAN66957.1; -  
DR TIGR; PP1334; -  
DR GO; GO:0016020; C:membrane; IEA.  
DR GO; GO:0008963; F:phospho-N-acetylmuramoyl-pentapeptide-trans. .; IEA.  
DR GO; GO:0006740; F:transferase activity; IEA.  
DR GO; GO:0006629; P:lipid metabolism; IEA.  
DR GO; GO:0009252; P:peptidoglycan biosynthesis; IEA.  
DR InterPro; IPR000715; Glyco. trans 4.  
DR InterPro; IPR003524; PNACPept trans.  
DR Pfam; PF00953; Glycos transf\_4\_1.  
DR PROSITE; PS01347; MRAY\_1; 1.

DR PROSITE; PS01348; MRAY\_2; 1.  
KW Transferase; Complete proteome.  
SQ SEQUENCE 360 AA; 39307 MW; 91B911C3B4118D56 CRC64;  
  
Query Match 91.9%; Score 1689; DB 16; Length 360;  
Best Local Similarity 90.3%; Pred. No. 1.5e-114;  
Matches 325; Conservative 18; Mismatches 17; Indels 0; Gaps 0;  
  
Qy 1 MLLLAAYLOQFYKGFVQYLTGRLGILSVLTALSLSLWLGPMWIRTLQIPQIGQAVRND 60  
Db 1 MLLLAAYLOQFYKGFVQYLTGRLGILSVLTALSLSLWLGPMWIRTLQIPQIGQAVRND 60  
  
Qy 61 GPQSHLSKKGPTMGGAIIPTAIAISTLLWADLSNRYVWVLTLLFGAIGWVDDYRKV 120  
Db 61 GPQSHLSKKGPTMGGAIIPTAIAISTLLWADLSNRYVWVLTLLFGAIGWVDDYRKV 120  
  
Qy 121 IEKNSRGLPSRWKYPWQSVFGIGAAVFLYMTAETPIETTLIVPMLKSVIEIQIGFFVLT 180  
Db 121 IEKNSRGLPSRWKYPWQSVFGIGAAVFLYMTAETPIETTLIVPMLKSVIEIQIGFFVLT 180  
  
Qy 181 YFVIIVGSSNAVLTDLGLAIMPVTVAGALGIFCYLSGNVKAFAEYLLIPNVPAGAGELI 240  
Db 181 YFVIIVGSSNAVLTDLGLAIMPVTVAGALGIFCYLSGNVKAFAEYLLIPNVPAGAGELI 240  
  
Qy 241 VFCAALVAGAGLGFNFNTYPAQVFMGVDGALGALGTIAIVRQEIIVLFIMGVFVME 300  
Db 241 VFCAALVAGAGLGFNFNTYPAQVFMGVDGALGALGTIAIVRQEIIVLFIMGVFVME 300  
  
Qy 301 TLSVMIQVASFKLTKRRVFRMAPIHFFELKMGWDPDRVIVRFFWITVILVIGLATLKL 360  
Db 301 TLSVMIQVASFKLTKRRVFRMAPIHFFELKMGWDPDRVIVRFFWITVILVIGLATLKL 360  
  
RESULT 2  
Q87WY2 PRELIMINARY; PRT; 360 AA.  
ID Q87WY2  
AC Q87WY2;  
DT 01-JUN-2003 (TremBLrel. 24, Created)  
DT 01-JUN-2003 (TremBLrel. 24, Last sequence update)  
DT 01-OCT-2003 (TremBLrel. 25, Last annotation update)  
DE Phospho-N-acetylmuramoyl-pentapeptide- transferase.  
GN MRAY OR PSPT04411.  
OS Pseudomonas syringae (pv. tomato).  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;  
OC Pseudomonadaceae; Pseudomonas.  
OX NCBI\_TaxID=3323;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=DC3000;  
RA Buell R., Joardar V., Khouri H., Fedorova N., Tran B., Russell D.,  
RA Berry K., Utterback T., Van Aken S., Feldblyum T., Gwinn M.,  
RA Dodson R., DeBoy R., Durkin A., Kolonay J., Madupu R., Daugherty S.,  
RA Brinkac L., Beanan M., Haft D., Selengut J., Nelson W., Davidsen T.,  
RA White O., Fraser C., Collier A.;  
RT "Complete sequence of Pseudomonas syringae."  
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AEO16871; AAC057860.1; -.  
DR TIGR; PSPT04411; -.  
DR GO; GO:0016020; C:membrane; IEA.  
DR GO; GO:0008963; P:phospho-N-acetylmuramoyl-pentapeptide-trans. .; IEA.  
DR GO; GO:0016740; P:transferase activity; IEA.  
DR GO; GO:0008629; P:lipid metabolism; IEA.  
DR GO; GO:0009252; P:peptidoglycan biosynthesis; IEA.  
DR InterPro; IPR000715; Glyco.trans.4.  
DR InterPro; IPR003524; PNAcPpept.trans.  
DR Pfam; PF00953; Glycos.transf.4; 1.  
DR PROSITE; PS01347; MRAY\_1; 1.  
DR PROSITE; PS01348; MRAY\_2; 1.  
KW Transferase; Complete proteome.  
SQ SEQUENCE 360 AA; 39484 MW; CF79F93DC0A8CD66 CRC64;  
  
Query Match 90.3%; Score 1658; DB 16; Length 360;  
Best Local Similarity 89.7%; Pred. No. 2.6e-112;  
Matches 325; Conservative 18; Mismatches 17; Indels 0; Gaps 0;

Matches 323; Conservative 17; Mismatches 20; Indels 0; Gaps 0;  
  
Qy 1 MLLLAAYLOQFYKGFVQYLTGRLGILSVLTALSLSLWLGPMWIRTLQIPQIGQAVRND 60  
Db 1 MLLLAAYLOQFYKGFVQYLTGRLGILSVLTALSLSLWLGPMWIRTLQIPQIGQAVRND 60  
  
Qy 61 GPQSHLSKKGPTMGGAIIPTAIAISTLLWADLSNRYVWVLTLLFGAIGWVDDYRKV 120  
Db 61 GPQSHLSKKGPTMGGAIIPTAIAISTLLWADLSNRYVWVLTLLFGAIGWVDDYRKV 120  
  
Qy 121 IEKNSRGLPSRWKYPWQSVFGIGAAVFLYMTAETPIETTLIVPMLKSVIEIQIGFFVLT 180  
Db 121 IEKNSRGLPSRWKYPWQSVFGIGAAVFLYMTAETPIETTLIVPMLKSVIEIQIGFFVLT 180  
  
Qy 181 YFVIIVGSSNAVLTDLGLAIMPVTVAGALGIFCYLSGNVKAFAEYLLIPNVPAGAGELI 240  
Db 181 YFVIIVGSSNAVLTDLGLAIMPVTVAGALGIFCYLSGNVKAFAEYLLIPNVPAGAGELI 240  
  
Qy 241 VFCAALVAGAGLGFNFNTYPAQVFMGVDGALGALGTIAIVRQEIIVLFIMGVFVME 300  
Db 241 VFCAALVAGAGLGFNFNTYPAQVFMGVDGALGALGTIAIVRQEIIVLFIMGVFVME 300  
  
Qy 301 TLSVMIQVASFKLTKRRVFRMAPIHFFELKMGWDPDRVIVRFFWITVILVIGLATLKL 360  
Db 301 TLSVMIQVASFKLTKRRVFRMAPIHFFELKMGWDPDRVIVRFFWITVILVIGLATLKL 360  
  
RESULT 3  
Q82VS6 PRELIMINARY; PRT; 361 AA.  
ID Q82VS6  
AC Q82VS6;  
DT 01-JUN-2003 (TremBLrel. 24, Created)  
DT 01-JUN-2003 (TremBLrel. 24, Last sequence update)  
DT 01-OCT-2003 (TremBLrel. 25, Last annotation update)  
DE Mray; phospho-N-acetylmuramoyl-pentapeptide-transferase (EC 2.7.8.13).  
GN MRAY OR NE0988.  
OS Nitrosomonas europaea.  
OC Bacteria; Proteobacteria; Betaproteobacteria; Nitrosomonadales;  
OC Nitrosomonadaceae; Nitrosomonas.  
OX NCBI\_TaxID=915;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=ATCC 19718 / IFO 14298;  
RX MEDLINE=22586410; PubMed=12700255;  
RA Chain P., Lamerdin J.E., Larimer F.W., Regala W., Lao V., Land M.,  
RA Hauser L., Hooper A.B., Klotz M.G., Norton J., Savavedra-Soto L.A.,  
RA Arciero D.M., Hommes N.G., Whittaker M.M., Arp D.J.;  
RT "Complete genome sequence of the ammonia-oxidizing bacterium and  
obligate chemolithoautotroph Nitrosomonas europaea."  
RL J. Bacteriol. 185:2759-2773(2003).  
DR EMBL; BX321859; CAD84899.1; -.  
DR GO; GO:0016020; C:membrane; IEA.  
DR GO; GO:0008963; P:phospho-N-acetylmuramoyl-pentapeptide-trans. .; IEA.  
DR GO; GO:0016740; P:transferase activity; IEA.  
DR GO; GO:0008629; P:lipid metabolism; IEA.  
DR GO; GO:0009252; P:peptidoglycan biosynthesis; IEA.  
DR InterPro; IPR000715; Glyco.trans.4.  
DR InterPro; IPR003524; PNAcPpept.trans.  
DR Pfam; PF00953; Glycos.transf.4; 1.  
DR TIGR; PS01348; MRAY\_2; 1.  
DR PROSITE; PS01347; MRAY\_1; 1.  
DR PROSITE; PS01348; MRAY\_2; 1.  
KW Transferase; Complete proteome.  
SQ SEQUENCE 361 AA; 39264 MW; ACSBF6C9523B2DD9 CRC64;  
  
Query Match 71.6%; Score 1314.5; DB 16; Length 361;  
Best Local Similarity 69.8%; Pred. No. 1.9e-87;  
Matches 252; Conservative 50; Mismatches 58; Indels 1; Gaps 1;





```
Db 241 INCTAIVAGLGLFYNTYPAQVFMGDSLSGLGALGVIAVLVRQELLLVWGGVFVE 300
QY 301 TLSVMIQVASFKLIGRRVFRMAPIHHHFKLKPWPFRVIVREWIITVLVIGLATLKL 360
Db 301 ALSVLQVGSYKLRQKRFMAPIHHHFKLKPWPFRVIVREWIITVLVIGLATLKL 360

RESULT 8
QWFR9 PRELIMINARY; PRT; 377 AA.
AC QWFR9;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Phospho-N-acetylmuramoyl-pentapeptide-transferase se (EC 2.7.8.13).
GN MRAY OR MURX OR B4201.
OS Bordetella bronchiseptica (Alcaligenes bronchisepticus).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Alcaligenaceae; Bordetella.
OX NCBI_TaxID=518;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RB50 / ATCC BAA-588;
RX MEDLINE=22827954; PubMed=12910271;
RA Parkhill J., Sebahia M., Preston A., Murphy L.D., Mungall K.L.,
RA Harris D.E., Holden M.T.G., Churcher C.M., Harris B., Quail M.A.,
RA Cerdano-Tarraga A.M., Temple L., James K., Basham D., Cherevach I.,
RA Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
RA Feltwell T., Goble A., Hamlin N., Hauser H., Holroyd S., Jagels K.,
RA Leather S., Moulé S., Norbertczak H., O'Neill S., Ormond D., Price C.,
RA Rabinowitsch E., Rutter S., Sanders M., Saunders D., Seeger K.,
RA Sharp S., Simmonds M., Skelton J., Squares R., Stevens K.,
RA Unwin L., Whitehead S., Barrell B.G., Maskell D.J.;
RT "Comparative analysis of the genome sequences of Bordetella pertussis,
RT Bordetella parapertussis and Bordetella bronchiseptica.";
RL Nat. Genet. 35:32-40(2003).
DR EMBL; BX640449; CAE34565.1; -.
KW Transferase; Complete proteome.
SQ SEQUENCE 377 AA; 41315 MW; 3440248BAD69BF1C CRC64;

Query Match 64.5%; Score 1184.5; DB 16; Length 377;
Best Local Similarity 62.8%; Pred. No. 5.3e-78;
Matches 236; Conservative 44; Mismatches 67; Indels 29; Gaps 3;

QY 14 KGFVFOYLTLRGILSVLTSLSLWLGPMWIRTQIPQIGQAVRNDGPOSHLSKKGTP 73
Db 2 RAIGVFYITLRAVLACATALLIGLVAGPRVIRRLTEKIGQAVRAYGPESHVKTGTP 61
QY 74 MCGALILTAISTLLWADLSNRYVMVVLVTLFGAIGWDDYRKVKIENRGLPSRWK 133
Db 62 MCGALILTAISTLLWADWTRNFVWVLLVTFGFGWIGMDDYRKVYRDPEGMPARQK 121
QY 134 YFQSVFGIGAAVFLYMTAETPIET-----TLIVPMLKSVEIQ 171
Db 122 FFWQATIGLVAAYLAFVAFSAFAPANTLWPLFKAWGSGFTMPLTRADLIVPFKSVSYP 181
QY 172 LGIF-FVLTIFYVIGSSNAVNLTGDLGLAIMPVWAGALGIFCYLSGNVKFAEYLLI 230
Db 182 LGVLGFALTWAVIGTSNAVNLTGDLGLAIMPVWGSALGIFAYVVGVDYSKYLLF 241
QY 231 PNVPAGELIVFCAALVAGLGLFWNTYPAQVFMGDSVAGALGALGATTAIVRQEVIL 290
Db 242 PYPGAAELMWLCAATIGGAGLAFNFNAYPAQVFMGDSVAGALGALGATTAIVRQEVIL 301
QY 291 FINGGVFMETLSVMIQVASFKLIT-----GRVFRMAPITHHFKLKPWPFRVIVREWI 344
Db 302 FINGGVFVETLSVMVQVTFWKYTKKYGQGRIRFMAPLHHHFEVGGWKEQVVVRFWI 361
QY 345 ITVLVIGLATLKL 360
Db 362 ITMMLVLVGLSTLKL 377

RESULT 9
QWVUQ PRELIMINARY; PRT; 377 AA.
AC QWVUQ;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Phospho-N-acetylmuramoyl-pentapeptide-transferase se (EC 2.7.8.13).
GN MRAY OR MURX OR BP3026.
OS Bordetella pertussis.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Alcaligenaceae; Bordetella.
OX NCBI_TaxID=520;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Tohana I / ATCC BAA-589 / NCTC 13251;
RX MEDLINE=22827954; PubMed=12910271;
RA Parkhill J., Sebahia M., Preston A., Murphy L.D., Mungall K.L.,
RA Harris D.E., Holden M.T.G., Churcher C.M., Harris B., Quail M.A.,
RA Cerdano-Tarraga A.M., Temple L., James K., Basham D., Cherevach I.,
RA Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
RA Feltwell T., Goble A., Hamlin N., Hauser H., Holroyd S., Jagels K.,
RA Leather S., Moulé S., Norbertczak H., O'Neill S., Ormond D., Price C.,
RA Rabinowitsch E., Rutter S., Sanders M., Saunders D., Seeger K.,
RA Sharp S., Simmonds M., Skelton J., Squares R., Stevens K.,
RA Unwin L., Whitehead S., Barrell B.G., Maskell D.J.;
RT "Comparative analysis of the genome sequences of Bordetella pertussis,
RT Bordetella parapertussis and Bordetella bronchiseptica.";
RL Nat. Genet. 35:32-40(2003).
DR EMBL; BX640420; CAE43297.1; -.
KW Transferase; Complete proteome.
SQ SEQUENCE 377 AA; 41315 MW; 3440248BAD69BF1C CRC64;

Query Match 64.5%; Score 1184.5; DB 16; Length 377;
Best Local Similarity 62.8%; Pred. No. 5.3e-78;
Matches 236; Conservative 44; Mismatches 67; Indels 29; Gaps 3;

QY 14 KGFVFOYLTLRGILSVLTSLSLWLGPMWIRTQIPQIGQAVRNDGPOSHLSKKGTP 73
Db 2 RAIGVFYITLRAVLACATALLIGLVAGPRVIRRLTEKIGQAVRAYGPESHVKTGTP 61
QY 74 MCGALILTAISTLLWADLSNRYVMVVLVTLFGAIGWDDYRKVKIENRGLPSRWK 133
Db 62 MCGALILTAISTLLWADWTRNFVWVLLVTFGFGWIGMDDYRKVYRDPEGMPARQK 121
QY 134 YFQSVFGIGAAVFLYMTAETPIET-----TLIVPMLKSVEIQ 171
Db 122 FFWQATIGLVAAYLAFVAFSAFAPANTLWPLFKAWGSGFTMPLTRADLIVPFKSVSYP 181
QY 172 LGIF-FVLTIFYVIGSSNAVNLTGDLGLAIMPVWAGALGIFCYLSGNVKFAEYLLI 230
Db 182 LGVLGFALTWAVIGTSNAVNLTGDLGLAIMPVWGSALGIFAYVVGVDYSKYLLF 241
QY 231 PNVPAGELIVFCAALVAGLGLFWNTYPAQVFMGDSVAGALGALGATTAIVRQEVIL 290
Db 242 PYPGAAELMWLCAATIGGAGLAFNFNAYPAQVFMGDSVAGALGALGATTAIVRQEVIL 301
QY 291 FINGGVFMETLSVMIQVASFKLIT-----GRVFRMAPITHHFKLKPWPFRVIVREWI 344
Db 302 FINGGVFVETLSVMVQVTFWKYTKKYGQGRIRFMAPLHHHFEVGGWKEQVVVRFWI 361
QY 345 ITVLVIGLATLKL 360
Db 362 ITMMLVLVGLSTLKL 377

RESULT 10
QW4B1 PRELIMINARY; PRT; 377 AA.
ID QW4B1
AC QW4B1;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
```

```

OC Bradyrhizobiaceae; Bradyrhizobium.
OX NCBI_TaxID=375;
RN [1];
RN SEQUENCE FROM N.A.
RC STRAIN=USDA 110;
RC MEDLINE=22484998; PubMed=12597275;
RA Kaneko T., Nakamura Y., Sato S., Minamisawa K., Uchiumi T.,
RA Sasanoto S., Watanabe A., Iidesawa K., Iriiguchi M., Kawashima K.,
RA Kohara M., Matsumoto M., Shimo S., Tsuruoka H., Wada T., Yamada M.,
RA Tabata S.;
RT "Complete genomic sequence of nitrogen-fixing symbiotic bacterium
RT Bradyrhizobium japonicum USDA110.";
RL DNA Res. 9:189-197(2002).
RL EMBL; AP005959; BAC51870.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0008363; F:phospho-N-acetylserine-pentapeptide-trans. .; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006529; P:lipid metabolism; IEA.
DR GO; GO:0009252; P:peptidoglycan biosynthesis; IEA.
DR InterPro; IPR000715; Glyco-trans_4.
DR InterPro; IPR003524; FNACrpept_trans.
DR Pfam; PF00953; Glycos_transf_4; 1.
DR PROSITE; PS01347; MRAY_1; 1.
DR PROSITE; PS01348; MRAY_2; 1.
DR TRANSFERASE; Complete Proteome.
KW TRANSFERASE; Complete Proteome.
SQ SEQUENCE 367 AA; 39570 MW; A176529AED38E3E4 CRC64;

Query Match 53.1%; Score 976; DB 16; Length 367;
Best Local Similarity 52.5%; Pred. No. 6.5e-63;
Matches 189; Conservative 106; Indels 8; Gaps 2;

Qy 9 LQOFYKGF-----VFQYLTLRGLSVLTALSLSLWLPWMRTLQIPQ-IGQAVRND 60
Db 8 LSNTPGFGAFRTPLNFRVITTRTGAAGVTGALFVLFPGFWIIDHLRIROKGGPIRTD 67
Qy 61 GPQSHLSKGTPTMGGLIITATAISLTLLWADLSNRVWVVLVTLFGAIGWDDYRKV 120
Db 68 GPQSHLAKGTPPTMGGLISLGLTVGLWANPLNPNVIVLAVTLGFGVGFYDDYLVK 127
Qy 121 IEKNSRGLPSRWKYFWQSVFGIGAAVFLYMTAETPIETTLIVPMKSVIEQLGIFFFVLT 180
Db 128 TKQTTTGFSGKRLLLIEAATLVACYALVRLNRDPASTALTIPFKDVLHFGWFFVFG 187
Qy 181 YFVIVGSSNAVNTDGLDGLAIPTVNVAGALGFCVLSGNVKFAEYLLIENVPGAGSLI 240
Db 188 AFVIVGAGNAVLTLDGLDGLAIVPMVLTATSFAMIVLAGNAVADYLOIKYVAGTGELA 247
Qy 241 VFCAALYVAGLGLFWNTYPAQVPMGDVGALGAALGTIAVIVRQSVLTFIMGGVFVNE 300
Db 248 VLCALLAGLGLFWFNAPPASIFMGDTGSLAGMLGAIAVAKHIEIVLAVIGLGFVLE 307
Qy 301 TLSVWIVQVAFKLTGRVFEMAPTHHHEFLKGWDPDVIVRFEWITITVLVILGLATLKL 360
Db 308 AVSVIVQVSPFLTKGKIFRMAPIHHEFELQGWTEPQIVIRFWIISVWMLAAGLSTLKL 367

RESULT 12
Q7VQJ0
ID Q7VQJ0 PRELIMINARY; PRT; 365 AA.
AC Q7VQJ0;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DE 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Phospho-N-acetylserine-pentapeptide transferase (EC 2.7.8.13).
GN MRAY OR BFL139.
OS Candidatus Blochmannia floridanus.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; ant endosymbionts; Candidatus Blochmannia.
OX NCBI_TaxID=203907;
RN [1];
RN SEQUENCE FROM N.A.
RP MEDLINE=22784745; PubMed=12886019;
RA Gil R., Silva F.J., Zientz E., Delmotte F., Gonzalez-Candelas F.,

```

RA Latorre A., Rausell C., Kamerbeek J., Gadau J., Hoelldobler B.,  
 RA van Ham R.C.H.J., Gross R., Moya A.,  
 RT "The genome sequence of Blochmannia floridanus: comparative analysis  
 of reduced genomes."  
 RL Proc. Natl. Acad. Sci. U.S.A. 100:9388-9393 (2003).  
 DR EMBL: BX248584; CAD83660.1; --  
 KW Transferase; Complete proteome.  
 SQ SEQUENCE 365 AA; 41504 MW; E7A2EEFPC10965B14 CRC64;

Query Match 45.2%; Score 830.5; DB 16; Length 365;  
 Best Local Similarity 43.4%; Pred. No. 2.2e-52;  
 Matches 155; Conservative 88; Mismatches 111; Indels 3; Gaps 2;

OY 1 MLLLAELYLOQFKG--FGVFOYLTLRGILSVLTSLTSLSLWTGPMWIRTLOIP-QIGQAV 57  
 DB 1 MLFWIKITSYFISSTLFVWMSVIFRGVGGFFSLFISIVIGNRIIIVWLKRLMLQTI 60  
 OY 58 RNDGPQSHLSKKGPTMGGALILTAISTLWADLSNRYVWVVLVTLFGAIGWDDY 117  
 DB 61 RIDGPQSHLKYGTPTMGGIIILISVTSVITWSDLSNIYIWFILFIFVMYGILGLVDDF 120  
 OY 118 RKVIEKNSRGLSPRWKYPWQSVFGIGAAVFLYMTAETPIETTLIVPMLKSVEIQIGFV 177  
 DB 121 LKIKGDNIGLTLNKLWQSIAMTLIVIMFINRVYVENOSGLEFLRNIVCKLUKIDM 180  
 OY 178 VLTIFYVIGSSNAVNLTDGLDGLAIMPTVMVAGALGIFCYLSGNVKAFAEYLLIPNVP 237  
 DB 181 ILAYFVIVTSSNVLSNLDGLGVIVPVLVSVGLAIVTVVGVNIYITSDLVIERVDCIK 240  
 OY 238 ELIVFCAALVAGLGLPWLNTTPAQVFGMDVGCALGALGATIAIVVQRIIVLPMGVF 297  
 DB 241 ELVVCASIIAGLGLFWNSYPSQIFMGDVGSLGIVGILVSHLLHQEYLLLLMGIF 300  
 OY 298 VNETLSVMIQVASFKLTKGRVFRMAPIHFFELKGNPDPRVIVRWIIVLVLGL 354  
 DB 301 VIESLSVIFQSVYFKFKRIFKMAPIHHFFELKGNPEKIVVRVIVSSILVLSI 357

RESULT 13  
 Q8KGD1 PRELIMINARY; PRT; 368 AA.  
 ID Q8KGD1  
 AC Q8KGD1  
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Phospho-N-acetylmuramoyl-pentapeptide- transferase.  
 GN MRAY OR CT0037.  
 OS Chlorobium tepidum.  
 OC Bacteria; Chlorobi; Chlorobia; Chlorobiales; Chlorobiaceae;  
 OC Chlorobium.  
 OX NCBI\_TaxID=1097;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=TLS / ATCC 49652 / DSM 12025;  
 RX MEDLINE=22103685; PubMed=12093901;  
 RA Eisen J.A., Nelson K.E., Paulsen I. T., Heidelberg J.F., Wu M.,  
 RA Dodson R.J., Deboy R. J., Gwinn M.L., Nelson W.C., Haft D.H.,  
 RA Hickey E.K., Peterson J.D., Durkin A.S., Kolonay J.B., Yang F.,  
 RA Holt I., Umayam L.A., Mason T., Brenner M., Shea T.P., Parksey D.,  
 RA Nierman W.C., Feldblyum T.V., Hansen C.L., Craven M.B., Radune D.,  
 RA Vamathevan J., Khouri H., White O., Gruber T.M., Ketchum K.A.,  
 RA Venter J.C., Tettelin H., Bryant D.A., Fraser C.M.;  
 RT "The complete genome sequence of Chlorobium tepidum TLS, a  
 photosynthetic, anaerobic, green-sulfur bacterium";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:9509-9514 (2002).  
 DR EMBL: AE012784; AAM71285.1; --  
 DR TIGR; CT0037; --  
 DR GO: 0016020; C:membrane; IEA.  
 DR GO: 0008963; P:phospho-N-acetylmuramoyl-pentapeptide-trans. . .; IEA.  
 DR GO: 0016740; F:transferase activity; IEA.  
 DR GO: 0006629; P:lipid metabolism; IEA.  
 DR GO: 0009252; P:peptidoglycan biosynthesis; IEA.  
 DR InterPro; IPR000715; Glyco\_trans\_4.

DR InterPro; IPR003524; PNACpsept\_trans.  
 DR Pfam; PF00953; Glycos\_transf\_4; 1.  
 DR TIGRFAMs; TIGR00445; mray; 1.  
 DR PROSITE; PS01347; MRAY 1; 1.  
 DR PROSITE; PS01348; MRAY 2; 1.  
 KW Transferase; Complete proteome.  
 SQ SEQUENCE 368 AA; 40769 MW; 70F16FA3A3DE4F6A CRC64;

Query Match 45.1%; Score 828; DB 16; Length 368;  
 Best Local Similarity 43.3%; Pred. No. 3.4e-52;  
 Matches 161; Conservative 78; Mismatches 117; Indels 16; Gaps 5;

OY 1 MLLLAELYLOQFKY--GFGVFOYLTLRGILSVLTSLTSLSLWGLPFWMTIRTQIQIGQAVR 58  
 DB 1 MLVYILRYINELYSLPCMGVIEVLTFRASAAITALLIIIFAGORFIRFLK-SKFVEPIK 59  
 OY 59 NDGPQSHLSKKGPTMGGALILTAISTLWADLSNRYVWVVLVTLFGAIGWDDYR 118  
 DB 60 BEAPPEHRKKDQVPTMGGLMIIFAIEVSAPLNAKIDDPHVLMILAVFWMLGILGIDDDYQ 119  
 OY 119 KVIEKNSRGLSPRWKYPWQSVFGIGAAVFLYMTAETPIETTLI---VPMKLSVEIQIGI 174  
 DB 120 KVLVKVKGGLAGHYKLIGQVTLGL---VIGFYTWNDPVSFVLLSDTTVPFFKLSVDYGI 176  
 OY 175 FFVLTIVFVIGSSNAVNLTDGLDGLAIMPTVMVAGALGIFCYLSGNVKAFAEYLLIPNVP 234  
 DB 177 FYIPVVIITAVSVNAVNLTDGLDGLAAGNAAIVTFALGVFAYLCGNVAVSGYLSIPFIS 236  
 OY 235 GAGELIVFCAALVAGLGLFWNTTPAQVFGMDVGCALGALGATIAIVVQRIIVLFI 294  
 DB 237 GAGELIVFCAALVAGLGLFWNTTPAQVFGMDVGCALGALGATIAIVVQRIIVLFI 296  
 OY 295 GVFMETLSVMIQVASFKLTKGRVFRMAPIHFFELKGNPDPRVIVRWIIVLFI 348  
 DB 297 GVFMETLSVMIQVASFKLTKGRVFRMAPIHFFELKGNPDPRVIVRWIIVLFI 356  
 OY 349 LVILGLATLKLIR 360  
 DB 357 LFLTSLMTLKLIR 368

RESULT 14  
 Q8F4J3 PRELIMINARY; PRT; 368 AA.  
 ID Q8F4J3  
 AC Q8F4J3  
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Phospho-N-acetylmuramoyl-pentapeptide- transferase (EC 2.7.8.13).  
 GN MRAY1 OR LA2048  
 OS Leptospira interrogans.  
 OC Bacteria; Spirochaetes; Spirochaetales; Leptospiraceae; Leptospira.  
 OX NCBI\_TaxID=173;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=56601 / Serogroup Icterohaemorrhagiae / Serovar lai;  
 RA Ren S.;  
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AE011377; AAN49247.1; --  
 DR GO: 0016020; C:membrane; IEA.  
 DR GO: 0008963; P:phospho-N-acetylmuramoyl-pentapeptide-trans. . .; IEA.  
 DR GO: 0016740; F:transferase activity; IEA.  
 DR GO: 0006629; P:lipid metabolism; IEA.  
 DR GO: 0009252; P:peptidoglycan biosynthesis; IEA.  
 DR InterPro; IPR000715; Glyco\_trans\_4.  
 DR Pfam; PF00953; Glycos\_transf\_4; 1.  
 DR TIGRFAMs; TIGR00445; mray; 1.  
 DR PROSITE; PS01347; MRAY 1; 1.  
 KW Transferase; Complete proteome.  
 SQ SEQUENCE 368 AA; 40644 MW; 03071CAD6DE50C80 CRC64;

Query Match 44.6%; Score 819; DB 16; Length 368;

Best Local Similarity 45.8%; Score 782.5; DB 16; Length 353;  
Matches 167; Conservative 66; Mismatches 120; Indels 12; Gaps 3;  
QY 8 YLQOFY----KGFVQYLTLCILSVLTALSLWLGPMWIRTLOIPQIGAVRNDGPO 63  
DB 4 YLDLYFNHLDLSRIFSVTRFALMAGLTSMLVTFWGHKIIDFLYGLKFRSVRDDGPK 63  
QY 64 SHLSKGTPTMGGALILTAISTLLWADLSNRYVWVLTLLFGAIGWDDYRKVIEK 123  
DB 64 SHEIKGTPTMGGLIIIGSLISVLLWNLKNPNVILLSVFSLSFVGFADDDYMSVKK 123  
QY 124 NSRGLPSRWKYFQWSVEFG-IGAAVFLYMTAETPIE-----TLLVPMKLSVEIOLGIF 175  
DB 124 IKGMARARTFISILISFICILFFYTTGTTGQTKISPOLQDLFFPFKIGBVIALGII 183  
QY 176 FVLTYYFVIYVSSNAVNLTLDGLALMPTVMVAGALGIFCYLSGNVKAFAEYLLINPVFG 235  
DB 184 AIPPSILVIGSSHAVNLTLDGLDGLATGTVLISVMTLGVATYFSGTPIVANYLNIPVLP 243  
QY 236 AGEILVFCALVAGLGLFNFNTYPAQVFMGDVGCALALGAALGTIAIVVROEIVLIMG 295  
DB 244 AHEYSVFLSALTALFGLFNFAHPAQVFMGDTGSLFLGATLGMIVILLKKEILLILGA 303  
QY 296 VFVMTLSVNIQVASFKLTKGRVFRMAPIHHELFELKGMPPRVIVRPMIITVILVILGLA 355  
DB 304 IFVSEALSVILOVGSFKLTGKRIKFAPLHHELFELGGLKETKIVIRPMIIVILVILSLS 363  
QY 356 TLKUR 360  
DB 364 TLKIQ 368

RESULT 15  
Q9P172 PRELIMINARY; PRT; 353 AA.  
AC Q9P172;  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DE Phospho-N-acetylmuramoyl-pentapeptide- transferase (EC 2.7.8.13).  
GN MRAY OR CJO433C.  
OS Campylobacter jejuni.  
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;  
OC Campylobacteraceae; Campylobacter.  
OX NCBI\_TaxID=197;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=NCCTC 11168;  
RX MEDLINE=20150912; PubMed=10698204;  
RA Parkhill J., Wren B.W., Mungall K., Ketley J.M., Churcher C.,  
RA Basham D., Chillingworth T., Davies R.M., Feltwell T., Holtroyd S.,  
RA Jagals K., Karlyshev A.V., Moule S., Pallen M.J., Penn C.W.,  
RA Quail M.A., Rajandream M.A., Rutherford K.M., van Vliet A.H.M.,  
RA Whitehead S., Barrett B.G.;  
RT "The genome sequence of the food-borne pathogen Campylobacter jejuni  
RT reveals hypervariable sequences."  
RL Nature 403:665-668(2000).  
DR EMBL; AL139075; CAB74269.1; -.  
DR PIR; G81387; G81387.  
DR GO; GO:0016020; C:membrane; IEA.  
DR GO; GO:0003824; E:catalytic activity; IEA.  
DR GO; GO:0008963; E:phospho-N-acetylmuramoyl-pentapeptide-trans. . .; IEA.  
DR GO; GO:0006629; F:lipid metabolism; IEA.  
DR GO; GO:0009252; F:peptidoglycan biosynthesis; IEA.  
DR InterPro; IPR000715; Glyco.trans.4.  
DR InterPro; IPR003524; PNAcPpept.trans.  
DR Pfam; PF00953; Glycos.transf.4; 1.  
DR TIGRFAMs; TIGR00445; mray\_1; 1.  
DR PROSITE; PS01347; MRAY\_1; 1.  
DR PROSITE; PS01348; MRAY\_2; 1.  
KW Complete proteome.  
SQ SEQUENCE 353 AA; 39239 MW; 8EBBCFC4F1061C95 CRC64;

Query Match 42.6%; Score 782.5; DB 16; Length 353;  
Best Local Similarity 44.8%; Pred. No. 6.4e-49;  
Matches 155; Conservative 70; Mismatches 118; Indels 3; Gaps 2;  
QY 16 FGVQYLTLCILSVLTALSLWLGPMWIRTLOIPQIGAVRNDGPOSHLSKGTPTMG 75  
DB 10 YAFFYISVRAGFAFFIALCLSLFEMPKFITWAKAKNASQPIIYEAETHKTKCHTIMG 69  
QY 76 GALILTAISTLLWADLSNRYVWVLTLLFGAIGWDDYRKVIEK-NSRGLPSRWKY 134  
DB 70 GLIFISSAVIASLFCIKRFDNIFALSALLCLTLFLCLIGLIDDLGKLVKDNHSGLSPRMKL 129  
QY 135 FQWSVFGIGAAVFLYMTAETPIETTLIVPMKLSVEIOLGIFVFLVTVFVIVGSSNAVNLT 194  
DB 130 LAQIIAGLICILPLFYFSE--LSTELTFPFYKHPFLFDMEIFAIFAWILVLISSNAVNLT 187  
QY 195 DGLDGLAIMPTVMVAGALGIFCYLSGNVKAFAEYLLINPVGAGELIVFCALVAGLGL 254  
DB 188 DGLDGLATVPGIFSLSTLIGIFLYLSGNLYSEYLLPKIQGLGEVWIIICALIGALMGFL 247  
QY 255 WNTYPAQVFMGDVGCALALGAALGTIAIVVROEIVLIMGVFMETLSVMIQVASEKT 314  
DB 248 WYNCYPAQVFMGDSGLGFGIFGLAVISKNEILLILLIGFVLETVSVTLQVGSFKIF 307  
QY 315 GRVFRMAPIHHELFELKGMPPRVIVRPMIITVILVILGLATLKUR 360  
DB 308 NKRVFKMAPIHHELFELKGMPPRVIVRPMIITVILVILGLATLKUR 353

Search completed: May 7, 2004, 08:40:32  
Job time : 58 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2004, 08:37:37 ; Search time 18 Seconds  
(without alignments)  
1041.403 Million cell updates/sec

Title: US-10-089-787-2

Perfect score: 1837

Sequence: 1 MLLLLAEVLQFYKFGVQ.....RFWIIIVLVIGLATIKLR 360

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_42:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1824	99.3	360	1	MRAY_PSEAE
2	1689	91.9	360	1	MRAY_PSEBK
3	1858	90.3	360	1	MRAY_PSESM
4	1334.5	72.5	361	1	MRAY_XANAC
5	1327.5	72.3	361	1	MRAY_XANCP
6	1312	71.4	360	1	MRAY_SHEVI
7	1306	71.1	360	1	MRAY_SHEON
8	1296.5	70.6	361	1	MRAY_XYLFA
9	1293	70.4	360	1	MRAY_VIBPA
10	1292.5	70.4	361	1	MRAY_VIBFT
11	1292	70.3	360	1	MRAY_VIBCH
12	1291	70.3	360	1	MRAY_VIBVU
13	1274	69.4	360	1	MRAY_PASMU
14	1268	69.0	360	1	MRAY_ECOL6
15	1268	69.0	360	1	MRAY_ECOL1
16	1265	68.9	360	1	MRAY_YERPE
17	1264	68.8	360	1	MRAY_SALTI
18	1261	68.6	360	1	MRAY_SALTY
19	1216	66.2	360	1	MRAY_HAEIN
20	1186	64.6	360	1	MRAY_HAEDU
21	1135	61.8	360	1	MRAY_WIGBR
22	1100	59.9	360	1	MRAY_NEIMB
23	1099	59.8	360	1	MRAY_NEIMA
24	945	51.4	360	1	MRAY_BRUME
25	939	51.1	360	1	MRAY_RHILO
26	931	50.7	366	1	MRAY_RHIME
27	917	49.9	366	1	MRAY_AGRTS
28	915.5	49.8	357	1	MRAY_BUCAI
29	898.5	48.9	361	1	MRAY_RICMO
30	894	48.7	356	1	MRAY_ZYMMO
31	893.5	48.6	361	1	MRAY_RICRI
32	890.5	48.5	361	1	MRAY_RICCN
33	888.5	48.4	361	1	MRAY_RICTY

#### ALIGNMENTS

```

RESULT 1
MRAY_PSEAE
ID MRAY_PSEAE STANDARD; PRT; 360 AA.
AC Q9HVZ8; Q9EY47;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Phospho-N-acetylmuramoyl-pentapeptide-transferase (EC 2.7.8.13) (UDP-
DE MurNAC-pentapeptide phosphotransferase).
GN MRAY OR PA4415.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21178826; PubMed=11281713;
RA Azzolina B.A., Yuan X., Anderson M.S., El-Sherbeini M.;
RT "The cell wall and cell division gene cluster in the Mra operon of
RT Pseudomonas aeruginosa: cloning, production, and purification of
RT active enzymes.";
RL Protein Expr. Purif. 21:393-400(2001).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=ATCC 15692 / PA01;
MEDLINE=20437337; PubMed=10984043;
RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,
RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA Garber R.L., Goltry L., Tolentino E., Westbrock-Wadman S., Yuan Y.,
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reizer J., Saier M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
RT opportunistic pathogen.";
RL Nature 406:959-964(2000).
CC -!- FUNCTION: First step of the lipid cycle reactions in the
CC biosynthesis of the cell wall peptidoglycan (By similarity).
CC -!- CATALYTIC ACTIVITY: UDPMur2Ac(cyl-L-Ala-gamma-D-Glu-L-Lys-D-Ala-D-
CC Ala) + undecaprenyl phosphate = UMP + Mur2Ac(cyl-L-Ala-gamma-D-
CC Glu-L-Lys-D-Ala-D-Ala)-diphosphoundecaprenol.
CC -!- PATHWAY: Peptidoglycan biosynthesis
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
CC -!- SIMILARITY: Belongs to the glycosyltransferase family 4. Mray
CC subfamily.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; AY008276; AAC22121.1; -
CC EMBL; AB004856; AACG7803.1; -
CC PIR; H83094; H83094.

```

Q9zcw0 rickettsia  
P59436 buchnera ap  
O66465 aquifex aeo  
Q8rdq0 fusobacteri  
Q44776 borrelia bu  
Q9zly1 helicobacte  
O25235 helicobacte  
Q9k9e6 bacillus ha  
Q07322 staphylococ  
Q8nx36 staphylococ  
Q03521 bacillus su  
Q8cpk7 staphylococ

34 887.5 48.3 361 1 MRAY\_RICPR  
35 875 47.6 340 1 MRAY\_BUCBP  
36 853.5 46.5 359 1 MRAY\_AQUAE  
37 819 44.6 361 1 MRAY\_FUSNN  
38 812.5 44.2 351 1 MRAY\_BORBU  
39 709.5 38.6 353 1 MRAY\_HELPJ  
40 708.5 38.6 353 1 MRAY\_HELPY  
41 671 36.5 325 1 MRAY\_BACHD  
42 652 35.5 321 1 MRAY\_STAAM  
43 652 35.5 321 1 MRAY\_STAAS  
44 651 35.4 324 1 MRAY\_BACSU  
45 639 34.8 321 1 MRAY\_STAEP



Db 301 TSVVIVQVASFKLTKGRVFRMAPIHHPHFKGWPPEPRVIVRFTWITVLVLGLATLKLK 360

RESULT 3

ID MRAY\_PESM STANDARD; PRT; 360 AA.

AC Q87WY2;

DT 15-MAR-2004 (Rel. 43, Created)

DT 15-MAR-2004 (Rel. 43, Last sequence update)

DT 15-MAR-2004 (Rel. 43, Last annotation update)

DE Phospho-N-acetylmuramoyl-pentapeptide-transferase (EC 2.7.8.13) (UDP-

DE MurNac-pentapeptide phosphotransferase).

GN MRAY OR PSPT04411.

OS Pseudomonas syringae (pv. tomato).

OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;

OC Pseudomonadaceae; Pseudomonas.

OX NCBI\_TaxID=323;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=DC3000;

RX MEDLINE=22834015; PubMed=12928499;

RA Buell C.R., Joardar V., Lindeberg M., Selengut J., Paulsen I.T.,

RA Winn M.L., Dodson R.J., Deboy R.T., Durkin A.S., Kolonay J.F.,

RA Madupu R., Dougherty S., Brinkac L., Beanan M.J., Haft D.H.,

RA Nelson W.C., Davidson T., Zafar N., Zhou L., Liu J., Yuan Q.,

RA Khouri H., Fedorova N., Tran B., Russell D., Berry K., Utterback T.,

RA Van Aken S.E., Feldblyum T.V., D'Ascenzo M., Deng W.-L., Ramos A.R.,

RA Alfano J.R., Cartinour S., Chatterjee A.K., Delaney T.P.,

RA Lazarowitz S.G., Martin G.B., Schneider D.J., Tang X., Bender C.L.,

RA White O., Fraser C.M., Collmer A.;

RT "The complete genome sequence of the Arabidopsis and tomato pathogen

RT Pseudomonas syringae pv. tomato DC3000.";

RL Proc. Natl. Acad. Sci. U.S.A. 100:10181-10186(2003).

CC -!- FUNCTION: First step of the lipid cycle reactions in the

CC biosynthesis of the cell wall peptidoglycan (By similarity).

CC -!- CATALYTIC ACTIVITY: UDPMur2Ac(oyl-L-Ala-gamma-D-Glu-L-Lys-D-Ala-D-

CC Ala) + undecaprenyl phosphate = UMP + Mur2Ac(oyl-L-Ala-gamma-D-

CC Glu-L-Lys-D-Ala-D-Ala)-diphosphoundecaprenol.

CC -!- PATHWAY: Peptidoglycan biosynthesis

CC -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).

CC -!- SIMILARITY: Belongs to the glycosyltransferase family 4. Mray

CC subfamily.

CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.ebi.ac.uk/ebis-sib.ch/announcement/>

CC or send an email to [license@ebi-sib.ch](mailto:license@ebi-sib.ch)).

CC

DR EMBL; AB016871; AAO57860.1; --

DR TIGR; PSPT04411; --

DR HAMAP; MF\_00038; -- 1.

DR InterPro; IPR000715; Glyco\_trans\_4.

DR Pfam; PF00953; Glycos\_transf\_4; 1.

DR PROSITE; PS01347; MRAY\_1; 1.

DR PROSITE; PS01348; MRAY\_2; 1.

KW Peptidoglycan synthesis; Cell division; Transferase; Transmembrane;

KW Complete proteome.

FT TRANSMEM 21 43 Potential.

FT TRANSMEM 71 90 Potential.

FT TRANSMEM 97 114 Potential.

FT TRANSMEM 134 156 Potential.

FT TRANSMEM 168 185 Potential.

FT TRANSMEM 200 222 Potential.

FT TRANSMEM 235 257 Potential.

FT TRANSMEM 262 284 Potential.

FT TRANSMEM 289 311 Potential.

FT TRANSMEM 338 357 Potential.

SQ SEQUENCE 360 AA; 39484 MW; CF79F93DC0A8CD66 CRC64;

Query Match 90.3%; Score 1658; DB 1; Length 360;

Best Local Similarity 89.7%; Pred. No. 1.2e-110;

Matches 323; Conservative 17; Mismatches 20; Indels 0; Gaps 0;

QY 1 MILLLAAYLQOQYKGFQYLTALGILSVLTALSLWLGPMWIRTLOIPGOAVRND 60

DB 1 MILLLAAYLQOQYKGFQYLTALGILSVLTALSLWLGPMWIRTLOIPGOAVRND 60

QY 61 GPQSHLSKSGTPTMGALITLTAIALSTLWADLSNRYVWVVLVVTLLFCAIGVDDYRKV 120

DB 61 GPQSHLSKSGTPTMGALITLTAIALSTLWADLSNRYVWVVLVVTLLFCAIGVDDYRKV 120

QY 121 IEKNSRGLPSRWKYFWQSVFGIAAFLVMTAETIETTLIVPMLKSVEIQIGIFVILT 180

DB 121 IEKNSRGLPSRWKYFWQSVFGIAAFLVMTAETIETTLIVPMLKSVEIQIGIFVILT 180

QY 181 YFVIVGSSNAVLTDGLDGLAIMPVTVAGALGIFCYLSGNVKAFLYLLIPNVPAGELI 240

DB 181 YFVIVGSSNAVLTDGLDGLAIMPVTVAGALGIFCYLSGNVKAFLYLLIPNVPAGELI 240

QY 241 VFCAALVAGLGLFMTNTYPAQVFGDVGALGALGTIAIVROEIVLFTMGVFNVE 300

DB 241 VFSGALIGAGLGLFMTNTYPAQVFGDVGALGALGTIAIVROEIVLFTMGVFNVE 300

QY 301 TLSVMTQVASFKLTKGRVFRMAPIHHPHFKGWPPEPRVIVRFTWITVLVLGLATLKLK 360

DB 301 TLSVMTQVASFKLTKGRVFRMAPIHHPHFKGWPPEPRVIVRFTWITVLVLGLATLKLK 360

RESULT 4

MRAY\_XANAC STANDARD; PRT; 361 AA.

AC Q8PPE0;

DT 28-FEB-2003 (Rel. 41, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Phospho-N-acetylmuramoyl-pentapeptide-transferase (EC 2.7.8.13) (UDP-

DE MurNac-pentapeptide phosphotransferase).

GN MRAY OR XAC0777.

OS Xanthomonas axonopodis (pv. citri).

OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;

OC Xanthomonadaceae; Xanthomonas.

OX NCBI\_TaxID=92829;

RN [1]

RN SEQUENCE FROM N.A.

RC STRAIN=306 / ATCC 13902 / XV 101;

RX MEDLINE=22022145; PubMed=12024217;

RA da Silva A.C.R., Ferro J.A., Rainach F.C., Farah C.S., Furlan L.R.,

RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,

RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,

RA Camarotte G., Canhaman F., Cardozo J., Chambergo F., Ciapina L.P.,

RA Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorri H.,

RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,

RA Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,

RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,

RA Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,

RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,

RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,

RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,

RA Spindola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,

RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,

RA Scubal J.C., Kitajima J.P.;

RT "Comparison of the genomes of two Xanthomonas pathogens with differing

RT host specificities."

RL Nature 417:459-463(2002).

CC -!- FUNCTION: First step of the lipid cycle reactions in the

CC biosynthesis of the cell wall peptidoglycan (By similarity).

CC -!- CATALYTIC ACTIVITY: UDPMur2Ac(oyl-L-Ala-gamma-D-Glu-L-Lys-D-Ala-D-

CC Ala) + undecaprenyl phosphate = UMP + Mur2Ac(oyl-L-Ala-gamma-D-

CC Glu-L-Lys-D-Ala-D-Ala)-diphosphoundecaprenol.

CC -!- PATHWAY: Peptidoglycan biosynthesis

CC -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).

CC -!- SIMILARITY: Belongs to the glycosyltransferase family 4. Mray

CC subfamily.

CC

```
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE011708; AM35665.1; -.
CC DR HAMAP; MF 00038; -.
CC DR InterPro; IPR000715; Glyco trans 4.
CC DR Pfam; PF00953; Glycos transf 4; 1.
CC DR TIGRFAMs; TIGR00445; Mray_1; 1.
CC DR PROSITE; PS01347; MRAY_1; 1.
CC DR PROSITE; PS01348; MRAY_2; 1.
CC KW Peptidoglycan synthesis; Cell division; Transferase; Transmembrane;
CC Complete proteome.
CC FT TRANSMEM 21 43 POTENTIAL.
CC FT TRANSMEM 71 90 POTENTIAL.
CC FT TRANSMEM 97 114 POTENTIAL.
CC FT TRANSMEM 134 152 POTENTIAL.
CC FT TRANSMEM 165 187 POTENTIAL.
CC FT TRANSMEM 202 224 POTENTIAL.
CC FT TRANSMEM 231 253 POTENTIAL.
CC FT TRANSMEM 263 285 POTENTIAL.
CC FT TRANSMEM 290 312 POTENTIAL.
CC FT TRANSMEM 339 358 POTENTIAL.
CC SQ SEQUENCE 361 AA; 39626 MW; BC4B2A5027348AB CRC64;

Query Match 72.3%; Score 1331.5; DB 1; Length 361;
Best Local Similarity 70.4%; Pred. No. 1.6e-87;
Matches 254; Conservative 43; Mismatches 63; Indels 1; Gaps 1;

Qy 1 MLLLLAEYLQOFYKGFYKGYLTLRGLTSLVLTSLSLWLGPMWIRTLPQIGQAVRND 60
Db 1 MLELARWLQQLSELPGLFNYLTFRGILAAITLALFSLWMPAVIRKLAQFGGQPIRD 60
Qy 61 GPQSHLSKGTPTMGGALITAIATISLLWADLSNRYVWVVLVTLFGAIGVDDYRKV 120
Db 61 GPQTHFSKAGTPTMGGSLITLTVLSVLGDLNRXVWLAVMVICFAGIWDWIKI 120
Qy 121 IEKNSRLGPRWKYFQWSVFGIGNAFLYMTATPTETTLIVPKLSVEIQL-GIFVVL 179
Db 121 VKRDPNGKRWKRYLLOSIFGLAAGFLYYTADVPAAITPIPMFKSIALPLAGVSFVVI 180
Qy 180 TYFVIVGSSNAVNLTDLGLDGLAIMPTVWAGALGIFCYLSGNKVFAYELIIPNVPGAGEL 239
Db 181 AYFWIVGFSNAVNLTDLGLDGLAIMPTVLVACALGVFAYAGSNVVFAYELKIPILPGAGEL 240
Qy 240 IVFCAALVAGLGLFLENTYPAQVPMGDVGALALGAALGTIAIVIVRQEVLFVFMGVFM 299
Db 241 IICSAIAGAGLGLFLENTYPAQVPMGDIGALSGLAVLGTIAIVIVRQEMVLVIMGVFVI 300
Qy 300 ETLSVMTQVASFKLTKRRVFRMAPIHFFHFKGPDPRVIVRWIITVILVILGLATKL 359
Db 301 ETLSVMTQVASFKLTKRRVFRMAPIHFFHFKGPDPRVIVRWIISVLVILGLATLKV 360
Qy 360 R 360
Db 361 R 361

RESULT 5
MRAY_XANCP STANDARD; PRT; 361 AA.
AC Q8PCK2;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Phospho-N-acetyl-muramoyl-pentapeptide-transferase (EC 2.7.8.13) (UDP-
DE MuraNac-pentapeptide phosphotransferase).
```

```
GN MRAY OR XCC0723.
OS Xanthomonas campestris (pv. campestris).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
OX NCBI_TaxID=340;
RN 1;
RP SEQUENCE FROM N.A.
RC STRAIN=ANCC 33913 / NCPPB 528;
RX MEDLINE=22022145; PubMed=1204217;
RA da Silva A.C.R., Ferro J.A., Reinach P.C., Farah C.S., Furlan L.R.,
RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,
RA Camarotte G., Cannavan F., Cardozo J., Chamberg F., Ciapina L.P.,
RA Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorri H.,
RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
RA Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
RA Locati E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.H.,
RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
RA Spinola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.P.,
RA Setubal J.C., Kitajima J.P.;
RT "Comparison of the genomes of two Xanthomonas pathogens with differing
RT host specificities.";
RL Nature 417:459-463(2002).
CC -!- FUNCTION: First step of the lipid cycle reactions in the
CC biosynthesis of the cell wall peptidoglycan (By similarity).
CC -!- CATALYTIC ACTIVITY: UDPMur2Ac(Oyl)-l-Ala-gamma-D-Glu-L-Lys-D-Ala-D-
CC Ala) + undecaprenyl phosphate = UMP + Mur2Ac(Oyl)-L-Ala-gamma-D-
CC Glu-L-Lys-D-Ala-D-Ala)-diphosphoundecaprenol.
CC -!- PATHWAY: Peptidoglycan biosynthesis.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
CC -!- SIMILARITY: Belongs to the glycosyltransferase family 4. Mray
CC subfamily.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE012171; AM40038.1; -.
CC DR HAMAP; MF 00038; -.
CC DR InterPro; IPR000715; Glyco trans 4.
CC DR InterPro; IPR003524; PNAcPept trans.
CC DR Pfam; PF00953; Glycos transf 4; 1.
CC DR TIGRFAMs; TIGR00445; Mray_1; 1.
CC DR PROSITE; PS01347; MRAY_1; 1.
CC DR PROSITE; PS01348; MRAY_2; 1.
CC KW Peptidoglycan synthesis; Cell division; Transferase; Transmembrane;
CC Complete proteome.
CC FT TRANSMEM 21 43 POTENTIAL.
CC FT TRANSMEM 71 90 POTENTIAL.
CC FT TRANSMEM 97 114 POTENTIAL.
CC FT TRANSMEM 134 152 POTENTIAL.
CC FT TRANSMEM 165 187 POTENTIAL.
CC FT TRANSMEM 202 224 POTENTIAL.
CC FT TRANSMEM 231 253 POTENTIAL.
CC FT TRANSMEM 263 285 POTENTIAL.
CC FT TRANSMEM 290 312 POTENTIAL.
CC FT TRANSMEM 339 358 POTENTIAL.
CC SQ SEQUENCE 361 AA; 39606 MW; PFS47C4E4FD21B8 CRC64;

Query Match 72.3%; Score 1327.5; DB 1; Length 361;
Best Local Similarity 69.8%; Pred. No. 3.1e-87;
Matches 252; Conservative 44; Mismatches 64; Indels 1; Gaps 1;

Qy 1 MLLLLAEYLQOFYKGFYKGYLTLRGLTSLVLTSLSLWLGPMWIRTLPQIGQAVRND 60
```

```

Db 1 MLLELRLWLQLESFLGLFNVLTRFGLAALTALFLSLMGMGPAVIRKLAQFKGQPIRQD 60
Qy 61 GPQSHLSKKGPTMGAGLILTAISTLLWADLSNRYVWVVLVTLTLLFGAIGWDDYRKV 120
Db 61 GPQSHLSKKGPTMGAGLILTAISTLLWADLSNRYVWVVLVTLTLLFGAIGWDDYRKV 120
Qy 121 IEKNSRGLPSRWKYFWQSVFGIGAAVFLYMTAETPIETTLIVPMLKSVIEIOL-GIFFVVL 179
Db 121 IEKNSRGLPSRWKYFWQSVFGIGAAVFLYMTAETPIETTLIVPMLKSVIEIOL-GIFFVVL 179
Qy 180 TVFVLVGSNAVNLTDGLDGLAINTVAVAGLGFYCYLSGNVFAEYLLIPNVPAGEL 239
Db 180 TVFVLVGSNAVNLTDGLDGLAINTVAVAGLGFYCYLSGNVFAEYLLIPNVPAGEL 239
Qy 240 IVFCAALVAGLGFYFWNTYPAQVPMGVDGALGALGALGTIAVIVROEIVLFGVGVFM 299
Db 240 IVFCAALVAGLGFYFWNTYPAQVPMGVDGALGALGALGTIAVIVROEIVLFGVGVFM 299
Qy 300 ETLNMIQVASFPLTGRVFRMAPIHHPFLKGPWDPDRVIVRFVITVILVIGLTLKL 359
Db 300 ETLNMIQVASFPLTGRVFRMAPIHHPFLKGPWDPDRVIVRFVITVILVIGLTLKL 359
Qy 360 R 360
Db 361 R 361

RESULT 6
MRAY_SHEV1
ID MRAY_SHEV1 STANDARD; PRT; 360 AA.
AC Q9FIN3;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE Phospho-N-acetylmuramoyl-pentapeptide-transferase (EC 2.7.8.13) (UDP-
MURNAc-pentapeptide phosphotransferase).
GN MRAY.
OS Shewanella violacea.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Alteromonadales;
OC Alteromonadales; Shewanella.
OX NCBI_TaxID=60217;
RN [1]
RP SEQUENCE FROM N.A.
RA Ishii A., Nakasone K., Sato T., Sugai M., Wachi M., Nagai K., Kato C.;
RT "Isolation and characterization of dcw gene cluster for cell division
RT and cell wall synthesis from a deep-sea piezophilic Shewanella
RT violacea.";
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: First step of the lipid cycle reactions in the
CC biosynthesis of the cell wall peptidoglycan (By similarity).
CC -!- CATALYTIC ACTIVITY: UDPMur2Ac(oyl-L-Ala-gamma-D-Glu-L-Lys-D-Ala-D-
CC Ala) + undecaprenyl phosphate = UMP + Mur2Ac(oyl-L-Ala-gamma-D-
CC Glu-L-Lys-D-Ala-D-Ala)-diphosphoundecaprenol.
CC -!- PATHWAY: Peptidoglycan biosynthesis.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
CC -!- SIMILARITY: Belongs to the glycosyltransferase family 4. Mray
CC subfamily.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AB052554; BAB19199.1; --
CC HAMAP; MF_00038; -- 1.
CC InterPro; IPR000715; Glyco.trans.4.
CC InterPro; IPR003524; PNACPept.trans.
CC Pfam; PF00953; Glycos.transf.4; 1.
CC TIGRFAMs; TIGR00445; mray; 1.
CC PROSITE; PS01347; MRAY_1; 1.

```

```

DR PROSITE; PS01348; MRAY_2; 1.
KW Peptidoglycan synthesis; Cell division; Transferase; Transmembrane.
FT TRANSMEM 20 43 POTENTIAL.
FT TRANSMEM 74 93 POTENTIAL.
FT TRANSMEM 97 114 POTENTIAL.
FT TRANSMEM 134 156 POTENTIAL.
FT TRANSMEM 171 193 POTENTIAL.
FT TRANSMEM 200 219 POTENTIAL.
FT TRANSMEM 234 256 POTENTIAL.
FT TRANSMEM 263 285 POTENTIAL.
FT TRANSMEM 289 311 POTENTIAL.
FT TRANSMEM 338 357 POTENTIAL.
SQ SEQUENCE 360 AA; 39777 MW; 53B2A2BA6F6B6FE4 CRC64;

Query Match 71.4%; Score 1312; DB 1; Length 360;
Best Local Similarity 69.2%; Pred. No. 3.9e-86;
Matches 249; Conservative 44; Mismatches 67; Indels 0; Gaps 0;

Qy 1 MLLLLAHYQQFYKGVFOYTLRCILSVLTSLSLWLGPMWIRTLQIPQIGQAVRND 60
Db 1 MLVYLAELYTQFYSGFNFSYVTFRAILGLMTALVFLWGWGPKMIRRLTLQIGQVVRSD 60
Qy 61 GPQSHLSKKGPTMGAGLILTAISTLLWADLSNRYVWVVLVTLTLLFGAIGWDDYRKV 120
Db 61 GPQSHLSKKGPTMGAGLILTAISTLLWADLSNRYVWVVLVTLTLLFGAIGWDDYRKV 120
Qy 121 IEKNSRGLPSRWKYFWQSVFGIGAAVFLYMTAETPIETTLIVPMLKSVIEIOLGIFPVVLT 180
Db 121 VRKDTKGLIARWKYILQSLAAIIAFYLYASADTVGETQLVVPFPFKDIMPQMGAPFVLA 180
Qy 181 YFVIVGSSNAVNLTDGLDGLAINTVAVAGLGFYCYLSGNVFAEYLLIPNVPAGELI 240
Db 181 YFTVIGSSNAVNLTDGLDGLAINTVAVAGLGFYCYLSGNVFAEYLLIPNVPAGELV 240
Qy 241 VFCALVAGLGFYFWNTYPAQVPMGVDGALGALGALGTIAVIVROEIVLFGVGVFVME 300
Db 241 IVCTAVGAGLGFYFWNTYPAQVPMGVDGSLGALGALGTIAVIVROEIVLFGVGVFVME 300
Qy 301 TLSVMIQVASFPLTGRVFRMAPIHHPFLKGPWDPDRVIVRFVITVILVIGLTLKL 360
Db 301 TVSVILQVSGYKLRGRIFRMAPIHHPFLKGPWDPDRVIVRFVITVILVIGLTLKL 360

RESULT 7
MRAY_SHEON
ID MRAY_SHEON STANDARD; PRT; 360 AA.
AC Q8E9F5;
DT 15-MAR-2004 (Rel. 43, Created)
DT 15-MAR-2004 (Rel. 43, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Phospho-N-acetylmuramoyl-pentapeptide-transferase (EC 2.7.8.13) (UDP-
MURNAc-pentapeptide phosphotransferase).
GN MRAY OR S04222.
OS Shewanella oneidensis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Alteromonadales;
OC Alteromonadales; Shewanella.
OX NCBI_TaxID=70863;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=MR-1;
RA MEDLINE=22297666; PubMed=12368813;
RA Heidelberg J.F., Paulsen I.T., Nelson K.E., Gaidos E.J., Nelson W.C.,
RA Read T.D., Eisen J.A., Seshadri R., Ward N., Methe B., Clayton R.A.,
RA Meyer T., Teapin A., Scott J., Beanan M., Brinkac L., Daugherty S.,
RA DeBoy R.T., Dodson R.J., Durkin A.S., Haft D.H., Kolonay J.F.,
RA Madupu R., Peterson J.D., Unayam L.A., White O., Wolf A.M.,
RA Vamathevan J., Weidman J., Imbraim M., Lee K., Berry K., Lee C.,
RA Mueller J., Khouri H., Gill J., Utterback T.R., McDonald L.A.,
RA Feldblyum T.V., Smith H.O., Venter J.C., Nealon K.H., Fraser C.M.;
RT "Genome sequence of the dissimilatory metal ion-reducing bacterium
RT Shewanella oneidensis.";
RL Nat. Biotechnol. 20:1118-1123 (2002).
CC -!- FUNCTION: First step of the lipid cycle reactions in the

```

CC biosynthesis of the cell wall peptidoglycan (By similarity).

CC -1- CATALYTIC ACTIVITY: UDPMur2Ac (Oyl-L-Ala-gamma-D-Glu-L-Lys-D-Ala-D-

CC Ala) + undecaprenyl phosphate = UMP + Mur2Ac (Oyl-L-Ala-gamma-D-

CC Glu-L-Lys-D-Ala-D-Ala)-diphosphoundecaprenol.

CC -1- PATHWAY: Peptidoglycan biosynthesis.

CC -1- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).

CC -1- SIMILARITY: Belongs to the glycosyltransferase family 4. Mray

CC subfamily.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

CC EMBL; AE015855; AAN57194.1; -

CC TIGR; SO4222; -

CC HAMAP; MF 00038; -; 1.

CC InterPro; IPR000715; Glyco trans 4.

CC InterPro; IPR003524; PNACPpt\_trans.

CC Pfam; PF00953; Glycos\_transf\_4; 1.

CC TIGRfams; TIGR00445; mray; 1.

CC PROSITE; PS01347; MRAY\_1; 1.

CC PROSITE; PS01348; MRAY\_2; 1.

CC Peptidoglycan synthesis; Cell division; Transferase; Transmembrane;

CC Complete proteome.

CC TRANSMEM 20 42 Potential.

CC TRANSMEM 74 93 Potential.

CC TRANSMEM 97 114 Potential.

CC TRANSMEM 134 153 Potential.

CC TRANSMEM 168 190 Potential.

CC TRANSMEM 197 219 Potential.

CC TRANSMEM 234 256 Potential.

CC TRANSMEM 263 285 Potential.

CC TRANSMEM 289 311 Potential.

CC TRANSMEM 338 357 Potential.

CC SEQUENCE 360 AA; CF447768C8F73217 CRC64;

CC

Query Match 71.1%; Score 1306; DB 1; Length 360;

Best Local Similarity 68.6%; Pred. No. 1e-85;

Matches 247; Conservative 47; Mismatches 66; Indels 0; Gaps 0;

Oy 1 MLLLLAEYLQFVKGFGVGYLTALRGLSVLTALSLSMLGPMWRTLQIPQIGQAVRND 60

Db 1 MLVYLAELYLTFHTGFNVFSYVTFRALIGLTLALMFLSWGPKLIERLQMQIGQVND 60

Oy 61 GPQSHLSKGGPTMGGLIITATAISTLLWADLSNRYVWVLLVVTLLFGAIGWVDYRKV 120

Db 61 GPESHFSKRGPTMGGLIITGAIFLSVLLWGLDLSRYVWVLLVVTLLFGAIGWVDYRKV 120

Oy 121 IEKNSRGLPGRWKYFMQSVFGIGAANFLVMTAETPIETTLIVPMKSVIEIQIGFFVLT 180

Db 121 VRKDTKGLARWKYILQSLAALITAFLLYTTAANPGETQLVWFFKDWMPQLGAVFVLA 180

Oy 181 YFVIVGSSNAVLTDGLDGLATPTWVACALGIFVCLSGNVKFAEYLLIPNVFAGELI 240

Db 181 YFTIVGSSNAVLTDGLDGLATPTWVAAAFALIALYLSHAQFANYLHPHLPGSGELV 240

Oy 241 VFCAALVAGAGLGFNFNTYPAQFMGVDGALALCAALGTATVAVROEIVLFMGVFWYNE 300

Db 241 IVCTAIVGAGLGFNFNTYPAQFMGVDGSLSLGAALGALAVLRQEILLVINGGVFWYNE 300

Oy 301 TLSVMIQVASKLTGRVFRMAPTHHFLKGHDPDRVIVFWIIVTIVLGLATLKLK 360

Db 301 TVSVILQVGSYKLRGQIFRMAPTHHHYELKGPPEVIVFWIIVTIVLGLATLKLK 360

RESULT 8

MEAY\_XYLFPA

ID\_MRAY\_XYLFPA

AC\_Q9PF83; STANDARD; PRT; 361 AA.

DT 28-FEB-2003 (Rel. 41, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Phospho-N-acetylmuramoyl-pentapeptide-transferase (EC 2.7.8.13) (UDP-

DE MurNac-pentapeptide phosphotransferase).

GN XRAY OR XF0795

OS Xylella fastidiosa.

OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;

OC Xanthomonadaceae; Xylella.

OX NCBI\_TaxID=2371;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=9a5C;

EX MEDLINE=20365717; PubMed=10910347;

RA Simpson A.J.G.; Reinach F.C.; Arruda P.; Abreu F.A.; Acencio M.

RA Alvares A.R.; Alves D.M.C.; Araya J.E.; Sata G.S.; Baptista C.S.

RA Barros M.H.; Bonaccorsi E.D.; Bordin S.; Bove J.M.; Briones M.R.S.

RA Bueno M.R.P.; Camargo A.A.; Camargo L.E.A.; Carraro D.M.; Carier H.

RA Colauto N.B.; Colombo C.; Costa F.F.; Costa M.C.R.; Costa-Neto C.M.

RA Coutinho L.L.; Cristofani M.; Dias-Neto E.; Docena C.; El-Dorry H.

RA Racincani A.P.; Ferreira A.J.S.; Ferreira V.C.A.; Ferro J.A.

RA Fraga J.S.; Franca S.C.; Franco M.C.; Frohme M.; Furlan L.R.

RA Garnier M.; Goldman G.H.; Goldman M.H.S.; Gomes S.L.; Gruber A.

RA Ho P.L.; Honeisel J.D.; Junqueira M.L.; Kemper E.L.; Kitajima J.P.

RA Krieger J.E.; Kuramae E.E.; Laigret F.; Lambais M.R.; Leite L.C.C.

RA Lemos E.G.M.; Lemos M.V.F.; Lopes S.A.; Lopes C.R.; Machado J.A.

RA Machado M.A.; Madeira A.M.B.N.; Madeira H.M.F.; Marino C.L.

RA Marques M.V.; Martins E.A.L.; Martins E.M.F.; Matsukuma A.Y.

RA Menck C.F.M.; Miracca E.C.; Miyaki C.Y.; Monteiro-Vitorello C.B.

RA Moon D.H.; Nagai M.A.; Nascimento A.L.T.O.; Netto L.E.S.

RA Nhani A.Jr.; Nobrega F.G.; Nunes L.R.; Oliveira M.A.

RA de Oliveira M.C.; de Oliveira R.C.; Palmieri D.A.; Paris A.

RA Peixoto B.R.; Pereira G.A.G.; Pereira H.A. Jr.; Pasquero J.B.

RA Quaggio R.B.; Roberto P.G.; Rodrigues V.; de Rosa A.J.M.

RA de Rosa V.E. Jr.; de Sa R.G.; Santelli R.V.; Sawasaki H.E.

RA da Silva A.C.R.; da Silva A.M.; da Silva F.R.; Silva W.A. Jr.

RA da Silveira J.F.; Silvestri M.L.Z.; Siqueira W.J.; de Souza A.A.

RA de Souza A.P.; Terenzi M.F.; Truffi D.; Tsai S.M.; Tshako M.H.

RA Vallada H.; Van Sluys M.A.; Verjovski-Almeida S.; Vettore A.L.

RA Zago M.A.; Zatz M.; Neidans J.; Setubal J.C.

RL "The genome sequence of the plant pathogen Xylella fastidiosa."

RL Nature 406:151-159 (2000).

CC -1- FUNCTION: First step of the lipid cycle reactions in the

CC biosynthesis of the cell wall peptidoglycan (By similarity).

CC -1- CATALYTIC ACTIVITY: UDPMur2Ac (Oyl-L-Ala-gamma-D-Glu-L-Lys-D-Ala-D-

CC Ala) + undecaprenyl phosphate = UMP + Mur2Ac (Oyl-L-Ala-gamma-D-

CC Glu-L-Lys-D-Ala-D-Ala)-diphosphoundecaprenol.

CC -1- PATHWAY: Peptidoglycan biosynthesis.

CC -1- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).

CC -1- SIMILARITY: Belongs to the glycosyltransferase family 4. Mray

CC subfamily.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

CC EMBL; AE003919; AAF83605.1; -

CC PIR; B82763; B82763.

CC HAMAP; MF 00038; -; 1.

CC InterPro; IPR000715; Glyco trans 4.

CC InterPro; IPR003524; PNACPpt\_trans.

CC Pfam; PF00953; Glycos\_transf\_4; 1.

CC TIGRfams; TIGR00445; mray; 1.

CC PROSITE; PS01347; MRAY\_1; 1.

CC PROSITE; PS01348; MRAY\_2; 1.

CC Peptidoglycan synthesis; Cell division; Transferase; Transmembrane;

CC Complete proteome.

CC TRANSMEM 21 43 POTENTIAL.

CC TRANSMEM 71 90 POTENTIAL.

FT TRANSMEM 97 114 POTENTIAL.  
FT TRANSMEM 139 161 POTENTIAL.  
FT TRANSMEM 168 187 POTENTIAL.  
FT TRANSMEM 202 224 POTENTIAL.  
FT TRANSMEM 237 259 POTENTIAL.  
FT TRANSMEM 263 285 POTENTIAL.  
FT TRANSMEM 290 312 POTENTIAL.  
FT TRANSMEM 339 358 POTENTIAL.  
SQ SEQUENCE 361 AA; 39694 MW; 83A40BDD66F990CF CRC64;  
  
Query Match 70.6%; Score 1296.5; DB 1; Length 361;  
Best Local Similarity 68.1%; Pred. No. 4.9e-85;  
Matches 246; Conservative 43; Mismatches 71; Indels 1; Gaps 1;  
  
QY 1 MLLLAELVQOQYKGFVFOYLTGLSLVLTALSLWLGPMWIRTQIQIGQAVRND 60  
DB 1 MFLFARWLQOQESLFGFLNYLTFSILAAULTALFSLWIGVLIQKLSQKQGPQRD 60  
QY 61 GPQSHLSKGTPTMGALILTAIATLSTLLWADLSNRYVWVLLVTLFGAIGWDDYRKV 120  
DB 61 GPKHFSKAGTPTMGSLIMTTLVSLVWGLRNRVYVWLVVWMLAFGAIGWDDYDKL 120  
QY 121 IEKNSRGLPSRWKYFWQSVFGIGAAVFLYMTAETPIETTLIVPMKSVIEIOL-GIFFVVL 179  
DB 121 ARDENGKLSRWKYLQIFGLAAGLFLYFTADVPAATVPIPFKSLALPLTSISFVAI 180  
QY 180 TYFVIVGSSNAVNLTDGLDGLAIMPVTMVAGALGIFCYLSGNVKAPEYLLIPNVPGAGEL 239  
DB 181 TYFVIVGSSNAVNLTDGLDGLAIMPVTMLVACALGVFAVAGNTLFPSSYKLTPTIPGAGDL 240  
QY 240 IVFCAALVAGLGLFWNTYPAQVFMGDVGVGALGALGTTIAIVVROSLVIFMGVYVME 299  
DB 241 IITCAIAGAGLGLFWNTYPAQVFMGDVGVGALGALGTTIAIVVROSLVIFMGVYVME 300  
QY 300 ETLVSMIOVASFKLTGRVFRMAPIHFFELKSGWDPDRVIVRFFWITVILVIGLIGLTKL 359  
DB 301 ETLVSMIOVASFKLTGRVFRMAPIHFFELKSGWDPDRVIVRFFWITVILVIGLIGLTKL 360  
QY 360 R 360  
DB 361 R 361  
  
RESULT 9  
MRAY\_VIBPA STANDARD; PRT; 360 AA.  
AC Q87SG7;  
DT 10-OCT-2003 (Rel. 42, Created)  
DT 10-OCT-2003 (Rel. 42, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Phospho-N-acetylmuramoyl-pentapeptide-transferase (EC 2.7.8.13) (UDP-MurNAc-pentapeptide phosphotransferase).  
GN MRAY OR VP0457.  
OS Vibrio parahaemolyticus.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;  
OC Vibrionaceae; Vibrio.  
OX NCBI\_TaxID=670;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=RIMD 2210633 / Serotype O3:K6;  
RX MEDLINE=22508454; PubMed=12620739;  
RA Makino K., Oshima K., Kurokawa K., Yokoyama K., Uda T., Tagomori K., Iijima Y., Najiima M., Nakano M., Yamashita A., Kubota Y., Kimura S., Yasunaga T., Honda T., Shinagawa H., Hattori M., Iida T.;  
RT "Genome sequence of Vibrio parahaemolyticus: a pathogenic mechanism distinct from that of V. cholerae";  
RL Lancet 361:743-749(2003).  
CC -!- FUNCTION: First step of the lipid cycle reactions in the biosynthesis of the cell wall peptidoglycan (By similarity).  
CC -!- CATALYTIC ACTIVITY: UDPMur2Ac(oyl-L-Ala-gamma-D-Glu-L-Lys-D-Ala-D-Ala) + undecaprenyl phosphate = UMP + Mur2Ac(oyl-L-Ala-gamma-D-Glu-L-Lys-D-Ala-D-Ala)-diphosphoundecaprenol.  
CC -!- PATHWAY: Peptidoglycan biosynthesis.

CC -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).  
CC -!- SIMILARITY: Belongs to the glycosyltransferase family 4. Mray subfamily.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; AP005074; BAC58720.1; -.  
CC HAMAP; MF 00038; -; 1.  
DR InterPro; IPR000715; Glyco trans 4.  
DR InterPro; IPR003524; PNACpept trans.  
DR Pfam; PF00953; Glycos transf\_4; 1.  
DR PROSITE; PS01347; MRAY\_1; 1.  
DR PROSITE; PS01348; MRAY\_2; 1.  
KW Peptidoglycan synthesis; Cell division; Transferase; Transmembrane; Complete proteome.  
FT TRANSMEM 23 45 Potential.  
FT TRANSMEM 74 93 Potential.  
FT TRANSMEM 97 114 Potential.  
FT TRANSMEM 134 151 Potential.  
FT TRANSMEM 171 193 Potential.  
FT TRANSMEM 200 219 Potential.  
FT TRANSMEM 234 256 Potential.  
FT TRANSMEM 263 285 Potential.  
FT TRANSMEM 289 311 Potential.  
FT TRANSMEM 338 357 Potential.  
SQ SEQUENCE 360 AA; 39943 MW; 4E032PA3F11C6B01 CRC64;  
  
Query Match 70.4%; Score 1293; DB 1; Length 360;  
Best Local Similarity 67.5%; Pred. No. 8.6e-85;  
Matches 243; Conservative 51; Mismatches 66; Indels 0; Gaps 0;  
  
QY 1 MLLLAELVQOQYKGFVFOYLTGLSLVLTALSLWLGPMWIRTQIQIGQAVRND 60  
DB 1 MLLLAELVQOQYKGFVFOYLTGLSLVLTALSLWLGPMWIRTQIQIGQAVRND 60  
QY 61 GPQSHLSKGTPTMGALILTAIATLSTLLWADLSNRYVWVLLVTLFGAIGWDDYRKV 120  
DB 61 GPQSHLSKGTPTMGALILTAIATLSTLLWADLSNRYVWVLLVTLFGAIGWDDYRKV 120  
QY 121 IEKNSRGLPSRWKYFWQSVFGIGAAVFLYMTAETPIETTLIVPMKSVIEIOLGIFVVL 180  
DB 121 VRKNTDGLIARWKYFWQSVLAFVFAFALYAGKDTAATQVLPFVKDVPQLGLMILT 180  
QY 181 YFVIVGSSNAVNLTDGLDGLAIMPVTMVAGALGIFCYLSGNVKAPEYLLIPNVPGAGELI 240  
DB 181 YFVIVGSSNAVNLTDGLDGLAIMPVTMVAGALGIFCYLSGNVKAPEYLLIPNVPGAGELI 240  
QY 241 VFCALVAGLGLFWNTYPAQVFMGDVGVGALGALGTTIAIVVROSLVIFMGVYVME 300  
DB 241 VFCALVAGLGLFWNTYPAQVFMGDVGVGALGALGTTIAIVVROSLVIFMGVYVME 300  
QY 301 TLSVMIOVASFKLTGRVFRMAPIHFFELKSGWDPDRVIVRFFWITVILVIGLIGLTKLR 360  
DB 301 TLSVMIOVASFKLTGRVFRMAPIHFFELKSGWDPDRVIVRFFWITVILVIGLIGLTKLR 360  
  
RESULT 10  
MRAY\_XYLFT STANDARD; PRT; 361 AA.  
AC Q87AF7;  
DT 10-OCT-2003 (Rel. 42, Created)  
DT 10-OCT-2003 (Rel. 42, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Phospho-N-acetylmuramoyl-pentapeptide-transferase (EC 2.7.8.13) (UDP-MurNAc-pentapeptide phosphotransferase).  
GN MRAY OR PD1868.  
OS Xylella fastidiosa (strain Temecula / ATCC 700954).



```

DR InterPro; IPR003524; PNAcPept_trans.
DR Pfam; PF00953; Glycos_transf_4; 1.
DR TIGRFAMs; TIGR00445; mray; 1.
DR PROSITE; PS01347; MRAY_1; 1.
DR PROSITE; PS01348; MRAY_2; 1.
KW Peptidoglycan synthesis; Cell division; Transferase; Transmembrane;
KW Complete proteome.
FT TRANSMEM 22 41 POTENTIAL.
FT TRANSMEM 74 93 POTENTIAL.
FT TRANSMEM 97 114 POTENTIAL.
FT TRANSMEM 134 151 POTENTIAL.
FT TRANSMEM 171 193 POTENTIAL.
FT TRANSMEM 200 222 POTENTIAL.
FT TRANSMEM 237 256 POTENTIAL.
FT TRANSMEM 263 285 POTENTIAL.
FT TRANSMEM 289 311 POTENTIAL.
FT TRANSMEM 338 357 POTENTIAL.
SQ SEQUENCE 360 AA; 39819 MW; 81727653AF899E53 CRC64;

Query Match 70.3%; Score 1292; DB 1; Length 360;
Best Local Similarity 66.9%; Pred. No. 1e-84;
Matches 241; Conservative 51; Mismatches 68; Indels 0; Gaps 0;

Qy 1 MLLLAELVQYKGFQVQYLTGRLSVLTALSLSLWLGPMWIRTLQIPQIGQAVRND 60
Db 1 MIILAEELQYFSPFLFEYLSFRAIVSILTALGISLWGMGRMTKRLQLQIGQVVRNE 60

Qy 61 GPQSHLSKGGPTMGALILTAIATSTLLWADLSNRYVWVVLVTLFAGLWDDYRKV 120
Db 61 GPESHFSRGPTMGVWVILAAITVLLWADLTNPYVWVAVLVLGAGVDDYRKV 120

Qy 121 IEKNSRGLPSRWKYFWQSGVFGIGAAVFLYMTAETPIETTLVPMKLSVEIQIGFFVLT 180
Db 121 VRKNTDGLIARWKYFWQSAIALVAFALYAGQDTAAQLVVPFKDVPQLGLMYILVT 180

Qy 181 YFVIVGSSNAVNLTDGLDGLAIMPVAVAGALGIFCYLSGNVKAELYLLIPNVPAGELI 240
Db 181 YFVIVGTSNAVNLTDGLDGLAIMPVAVAGFAIAWAGNPNFANYLHPIYFHSSELY 240

Qy 241 VFCAALVAGLGFLENFTYPAQVFMGDVGALAGALGTIAIVRQEIWLFMGCVFVME 300
Db 241 VVCTAMVAGLGFLENFTYPAQVFMGDVGALAGALGTIAIVRQEPFLVIMGVFVME 300

Qy 301 TLSVMIQVASFKLTGRVFRMAPIHFFELKGNWDPDRVIVRFTIIVLVLIGLATLKL 360
Db 301 TLSVLQVGSYKLRGQIRFMAPIHHYELKGNWPEPRVIVRFTIISVLVLIGLATLKV 360

RESULT 12
MRAY_VIEVU
ID MRAY_VIEVU STANDARD; PRT; 360 AA.
AC Q8DBK7;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Phospho-N-acetylmuramoyl-pentapeptide-transferase (EC 2.7.8.13) (UDP-
MURNAc-pentapeptide phosphotransferase).
GN MRAY OR VJ10561.
OS Vibrio vulnificus.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=672;
RN [1]_
RP SEQUENCE FROM N.A.
RC STRAIN=CMCP6;
RA Rhee J.H., Kim S.Y., Chung S.S., Kim J.J., Moon Y.H., Jeong H.,
RT Choy H.B.;
RL "Complete genome sequence of Vibrio vulnificus CMCP6.";
CC Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: First step of the lipid cycle reactions in the
CC biosynthesis of the cell wall peptidoglycan (By similarity).
CC -!- CATALYTIC ACTIVITY: UDPMurAc(oyl-L-Ala-gamma-D-Glu-L-Lys-D-Ala-D-
CC Ala) + undecaprenyl phosphate = UMP + Mur2Ac(oyl-L-Ala-gamma-D-

```

```

CC Glu-L-Lys-D-Ala-D-Ala)-diphosphoundecaprenol.
CC -!- PATHWAY: Peptidoglycan biosynthesis.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
CC -!- SIMILARITY: Belongs to the glycosyltransferase family 4. Mray
CC subfamily.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AEO16798; AAC09097.1; -.
CC HAMAP; MF_00038; 1.
DR InterPro; IPR000715; Glyco_trans_4.
DR InterPro; IPR003524; PNAcPept_trans.
DR Pfam; PF00953; Glycos_transf_4; 1.
DR TIGRFAMs; TIGR00445; mray; 1.
DR PROSITE; PS01347; MRAY_1; 1.
DR PROSITE; PS01348; MRAY_2; 1.
KW Peptidoglycan synthesis; Cell division; Transferase; Transmembrane;
KW Complete proteome.
FT TRANSMEM 15 37 POTENTIAL.
FT TRANSMEM 74 93 POTENTIAL.
FT TRANSMEM 97 114 POTENTIAL.
FT TRANSMEM 134 151 POTENTIAL.
FT TRANSMEM 171 193 POTENTIAL.
FT TRANSMEM 200 219 POTENTIAL.
FT TRANSMEM 234 256 POTENTIAL.
FT TRANSMEM 263 285 POTENTIAL.
FT TRANSMEM 289 311 POTENTIAL.
FT TRANSMEM 338 357 POTENTIAL.
SQ SEQUENCE 360 AA; 39850 MW; D087052F04DBA0B1 CRC64;

Query Match 70.3%; Score 1291; DB 1; Length 360;
Best Local Similarity 66.9%; Pred. No. 1.2e-84;
Matches 241; Conservative 55; Mismatches 64; Indels 0; Gaps 0;

Qy 1 MLLLAELVQYKGFQVQYLTGRLSVLTALSLSLWLGPMWIRTLQIPQIGQAVRND 60
Db 1 MIILAEELQYFSPFLFEYLSFRAIVSILTALGISLWGMGRMTKRLQLQIGQVVRNE 60

Qy 61 GPQSHLSKGGPTMGALILTAIATSTLLWADLSNRYVWVVLVTLFAGLWDDYRKV 120
Db 61 GPESHFSRGPTMGVWVILAAITVLLWADLTNPYVWVAVLVLGAGVDDYRKV 120

Qy 121 IEKNSRGLPSRWKYFWQSGVFGIGAAVFLYMTAETPIETTLVPMKLSVEIQIGFFVLT 180
Db 121 VRKNTDGLIARWKYFWQSAIALVAFALYAGQDTAAQLVVPFKDVPQLGLMYILVT 180

Qy 181 YFVIVGSSNAVNLTDGLDGLAIMPVAVAGALGIFCYLSGNVKAELYLLIPNVPAGELI 240
Db 181 YFVIVGTSNAVNLTDGLDGLAIMPVAVAGFAIAWAGNPNFANYLHPIYFHSSELY 240

Qy 241 VFCAALVAGLGFLENFTYPAQVFMGDVGALAGALGTIAIVRQEIWLFMGCVFVME 300
Db 241 VVCTAMVAGLGFLENFTYPAQVFMGDVGALAGALGTIAIVRQEPFLVIMGVFVME 300

Qy 301 TLSVMIQVASFKLTGRVFRMAPIHFFELKGNWDPDRVIVRFTIIVLVLIGLATLKL 360
Db 301 TLSVLQVGSYKLRGQIRFMAPIHHYELKGNWPEPRVIVRFTIISVLVLIGLATLKV 360

RESULT 13
MRAY_PASMU
ID MRAY_PASMU STANDARD; PRT; 360 AA.
AC P57816;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Phospho-N-acetylmuramoyl-pentapeptide-transferase (EC 2.7.8.13) (UDP-

```

DE MurNac-pentapeptide phosphotransferase).

GN MRAY OR PM0139.

OS Pasteurella multocida.

OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;

OC Pasteurellaceae; Pasteurella.

OX NCBI\_TaxID=747;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=PM70;

RX MEDLINE=21145866; PubMed=11248100;

RA May B.J., Zhang Q., Li L.L., Paustian M.L., Whittam T.S., Kapur V.;

RL "Complete genomic sequence of Pasteurella multocida PM70.";

RT Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).

CC -!- FUNCTION: First step of the lipid cycle reactions in the

CC biosynthesis of the cell wall peptidoglycan (By similarity).

CC -!- CATALYTIC ACTIVITY: UDPMur2Ac (Oyl-L-Ala-gamma-D-Glu-L-Lys-D-Ala-D-

CC Ala) + undecaprenyl phosphate = UMP + Mur2Ac (Oyl-L-Ala-gamma-D-

CC Glu-L-Lys-D-Ala-D-Ala)-diphosphoundecaprenol.

CC -!- PATHWAY: Peptidoglycan biosynthesis.

CC -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).

CC -!- SIMILARITY: Belongs to the glycosyltransferase family 4. Mray

CC subfamily.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC -----

CC

DR EMBL; AEO06048; AAK02223.1; -

DR HAMAP; MF 00038; -; 1.

DR InterPro; IPR000715; Glyco trans 4.

DR InterPro; IPR003524; PNACpept trans.

DR Pfam; PF00953; Glycos transf 4; 1.

DR TIGRfams; TIGR00445; mray\_1; 1.

DR PROSITE; PS01347; MRAY\_1; 1.

DR PROSITE; PS01348; MRAY\_2; 1.

DR Peptidoglycan synthesis; Cell division; Transferase; Transmembrane;

CC Complete proteome.

FT TRANSMEM 21 41 POTENTIAL.

FT TRANSMEM 73 93 POTENTIAL.

FT TRANSMEM 94 114 POTENTIAL.

FT TRANSMEM 132 152 POTENTIAL.

FT TRANSMEM 168 188 POTENTIAL.

FT TRANSMEM 200 220 POTENTIAL.

FT TRANSMEM 236 256 POTENTIAL.

FT TRANSMEM 263 283 POTENTIAL.

FT TRANSMEM 288 308 POTENTIAL.

FT TRANSMEM 338 358 POTENTIAL.

SQ SEQUENCE 360 AA; 40155 MW; B9BD1987FEDBD53 CRC64;

Query Match 69.4%; Score 1274; DB 1; Length 360;

Best Local Similarity 65.8%; Pred. No. 1.9e-83;

Matches 237; Conservative 57; Mismatches 66; Indels 0; Gaps 0;

QY 1 MLLLAELQOYKFGVGYLTALRGTLVSLTSLSLGPMWIRLQIPQGVAVND 60

DB 1 MLVWLGFLQOYSGFNVISITVTRAILALLTLLVSLWIGPKVIRLQLLKFQGVHRD 60

QY 61 GPQSHLKKGPTTGGGALITAIATISLLWADLSNRYVWVVLVTLFAGHWDDYRKV 120

DB 61 GPESHFSKRGPTTGGVWMLFAITVSTLLWANLANPYVWFSFLVLLGYGAIGFVDDYRKI 120

QY 121 IEKNSRGLPSNWKYFWQGVFGAIVFYMTAETPIETTLIVPMKSVETQLGFFVLT 180

DB 121 TRKNTDGLIARWKYFWLSVIALVAFGMYAIGKTDATRLVWPFKEIMPQLGLFIILT 180

QY 181 YFVIVGSSNANVLTDGLDGLAIMPTVMVAGALGFCFVLSGNVKAFAEYLLIPNVPGAGELI 240

DB 181 YFVIVGSSNANVLTDGLDGLAIMPTVMVAGAFALLIANATGNVNAFVILHPIYKFSALV 240

QY 241 VFCALYVAGLGLFWNTYPAQVFMGVDVGALALGALGTIATVIVRQEVILFIMGVFMVE 300

DB 241 VFCALYVAGLGLFWNTYPAQVFMGVDVGSIALGALGVVALVVRQEVILFIMGVFMVE 300

QY 301 TLSVMIOVASFKLTGRVFRMAPTHHFFELKSGWDPDRVIVRFTWITVLVIGATIKLR 360

DB 301 TLSVILQVSGYKLRKQIFRMAPTHHFFELKSGWDEPRVIVRFTWITVLVIGATIKLR 360

RESULT 14

MRAY\_ECOL6

ID MRAY\_ECOL6 STANDARD; PRT; 360 AA.

AC Q8X9Z0;

DT 28-FEB-2003 (Rel. 41, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Phospho-N-acetylmuramoyl-pentapeptide-transferase (EC 2.7.8.13) (UDP-

DE MurNac-pentapeptide phosphotransferase).

GN MRAY OR C0105 OR Z0097 OR ECS0091.

OS Escherichia coli O6, and

OS Escherichia coli O157:H7.

OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;

OC Enterobacteriaceae; Escherichia.

OX NCBI\_TaxID=217992, 83334;

RN [1]

RC SEQUENCE FROM N.A.

RC STRAIN=O6:H1 / CFT073 / ATCC 700928;

RX MEDLINE=2238234; PubMed=12471157;

RA Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,

RA Rasko D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud D.,

RA Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,

RA Mobley H.L.T., Donnenberg M.S., Blattner F.R.;

RT "Extensive mosaic structure revealed by the complete genome sequence

RT of uropathogenic Escherichia coli";

RC Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024(2002).

RN [2]

RC SEQUENCE FROM N.A.

RC STRAIN=O157:H7 / EDL933 / ATCC 700927;

RX MEDLINE=21074935; PubMed=11208551;

RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,

RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,

RA Posfai G., Hackett J., Link S., Boutin A., Shao Y., Miller L.,

RA Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamoculis K.,

RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,

RA Welch R.A., Blattner F.R.;

RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7";

RL Nature 409:529-533(2001).

RN [3]

RC SEQUENCE FROM N.A.

RC STRAIN=O157:H7 / RIMD 0509952;

RX MEDLINE=21156231; PubMed=11258796;

RA Hayashi T., Makino K., Nakayama K., Kurokawa K., Ishii K., Yokoyama K.,

RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,

RA Iida T., Takami H., Honda T., Sasaki C., Ogasawara N., Yasunaga T.,

RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;

RT "Complete genome sequence of enterohaemorrhagic Escherichia coli

RT O157:H7 and genomic comparison with a laboratory strain K-12";

RL DNA Res. 8:11-22(2001).

CC -!- FUNCTION: First step of the lipid cycle reactions in the

CC biosynthesis of the cell wall peptidoglycan (By similarity).

CC -!- CATALYTIC ACTIVITY: UDPMur2Ac (Oyl-L-Ala-gamma-D-Glu-L-Lys-D-Ala-D-

CC Ala) + undecaprenyl phosphate = UMP + Mur2Ac (Oyl-L-Ala-gamma-D-

CC Glu-L-Lys-D-Ala-D-Ala)-diphosphoundecaprenol.

CC -!- PATHWAY: Peptidoglycan biosynthesis.

CC -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).

CC -!- SIMILARITY: Belongs to the glycosyltransferase family 4. Mray

CC subfamily.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC -----

CC

entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

```
CC -----
CC DR EMBL; AE016755; AAN78603.1; -
CC DR EMBL; AE005185; AAG54391.1; -
CC DR EMBL; AF002550; BAB33514.1; -
CC DR PIR; C85491; C85491.
CC DR PIR; C90640; C90640.
CC DR HAMAP; MF_00038; -; 1.
CC DR InterPro; IPR000715; Glyco trans 4.
CC DR InterPro; IPR003524; PNAcPtept trans.
CC DR Pfam; PF00953; Glycos transf_4; 1.
CC DR TIGRfam; TIGR00445; mray; 1.
CC DR PROSITE; PS01347; MRAY_1; 1.
CC DR PROSITE; PS01348; MRAY_2; 1.
CC KW Peptidoglycan synthesis; Cell division; Transferase; Transmembrane;
CC Complete proteome.
FT DOMAIN 1 18 PERIPLASMIC (BY SIMILARITY).
FT TRANSMEM 19 45 BY SIMILARITY.
FT DOMAIN 46 76 CYTOPLASMIC (BY SIMILARITY).
FT TRANSMEM 77 90 BY SIMILARITY.
FT DOMAIN 91 96 PERIPLASMIC (BY SIMILARITY).
FT TRANSMEM 97 113 BY SIMILARITY.
FT DOMAIN 114 133 CYTOPLASMIC (BY SIMILARITY).
FT TRANSMEM 134 156 BY SIMILARITY.
FT DOMAIN 157 173 PERIPLASMIC (BY SIMILARITY).
FT TRANSMEM 174 188 BY SIMILARITY.
FT DOMAIN 189 199 CYTOPLASMIC (BY SIMILARITY).
FT TRANSMEM 200 220 BY SIMILARITY.
FT DOMAIN 221 238 PERIPLASMIC (BY SIMILARITY).
FT TRANSMEM 239 251 BY SIMILARITY.
FT DOMAIN 252 270 CYTOPLASMIC (BY SIMILARITY).
FT TRANSMEM 271 284 BY SIMILARITY.
FT DOMAIN 285 297 PERIPLASMIC (BY SIMILARITY).
FT TRANSMEM 298 300 BY SIMILARITY.
FT DOMAIN 300 341 CYTOPLASMIC (BY SIMILARITY).
FT TRANSMEM 342 357 BY SIMILARITY.
FT DOMAIN 358 360 PERIPLASMIC (BY SIMILARITY).
FT SEQUENCE 360 AA; 39903 MW; 922AB374BC9E8F2E CRC64;

Query Match
Best Local Similarity 66.4%; Pred. No. 5.le-83;
Matches 239; Conservative 51; Mismatches 70; Indels 0; Gaps 0;

OY 1 MULLIAEYQQYKFGVFOYLTLRGLSVLTALSLWLGPMWIRTQIPQIGAVRND 60
Db 1 MLVWLAHLVKYSGNFVESYITFRAIVSLTLTALFISLWMPRTAHQLKLSFGQVVRND 60

OY 61 GPQSHLSKSGTPTMGALILTALISTLWADLSNRYVWVLVTLFPAIGWDDYRKV 120
Db 61 GPESHFSKSGTPTMGGMILTAIVISVLWVPSNRYVWVLVTLFPAIGWDDYRKV 120

OY 121 IEKNSGLPSRWKYFWQSVFGGAFLVMTAETPIETTLIVPMLKSVETQIGFFVVL 180
Db 121 VRKDTAGLARKYFVMSVIALGVAPALYVGDTPATQVLPVFFKDVNPFQGLFYLLA 180

OY 181 YFVIVGSSNAVLTDGLDLAIPTVMVAGALGIFCYLSGNVKYFAEYLIPNVPAGBEL 240
Db 181 YFVIVGTGNAVLTDGLDLAIPTVFVAGGALVAVATGNMNFASYLHPIYLRHAGELV 240

OY 241 VFCAALVAGLGLFWNTYPAQVFMGDCVAGALGALGTIAIVIQEIVLFTMGGVFVME 300
Db 241 IVCTAVAGLGLFWNTYPAQVFMGDCVAGLGGALGIIAVLLRQEFLLVMGCVFVME 300

OY 301 TLSVMTQVASFKLITGRVFRMAPIHHLFELKGMPPDRVIVRVWITVITVILVIGLATLKL 360
Db 301 TLSVILQVGSFKLGRQIRFMAPIHHHVELKGMPPDRVIVRVWITVITVILVIGLATLKL 360

RESULT 15
ID MRAY_ECOLI
AC P15876;
STANDARD; PRT; 360 AA.
SPECIES=E.coli; STRAIN=K12;
FUNCTION.
```

```
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Phospho-N-acetylmuramoyl-pentapeptide-transferase (EC 2.7.8.13) (UDP-
DE MurNac-pentapeptide phosphotransferase).
GN MRAY OR MURX OR B0087 OR SF0084 OR S0086.
OS Escherichia coli; and
OS Shigella flexneri.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562, 623;
[1]
RN SEQUENCE FROM N.A.
RC SPECIES=E.coli; STRAIN=K12;
RX MEDLINE=90192099; PubMed=2179861;
RA Ikeda M., Wachi M., Ishino F., Matsuhashi M.;
RT "Nucleotide sequence involving mure and an open reading frame ORF-Y
RT spacing mure and ftsW in Escherichia coli.";
RN Nucleic Acids Res. 18:1058-1058(1990).
[2]
RN SEQUENCE FROM N.A.
RC SPECIES=E.coli; STRAIN=K12;
RX MEDLINE=92334977; PubMed=1630901;
RA Yura T., Mori H., Nagai H., Nagata T., Ishihama A., Fujita N.,
RA Isono K., Mizobuchi K., Nakata A.;
RT "Systematic sequencing of the Escherichia coli genome: analysis of
RT the 0-2.4 min region.";
RN Nucleic Acids Res. 20:3305-3308(1992).
[3]
RN SEQUENCE FROM N.A.
RC SPECIES=E.coli; STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner P.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RN Science 277:1453-1474(1997).
[4]
RN SEQUENCE FROM N.A.
RC SPECIES=S.flexneri; STRAIN=301 / Serotype 2a;
RX MEDLINE=2227406; PubMed=12384590;
RA Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,
RA Sun L., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,
RA Yang L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,
RA Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,
RA Yu J.;
RT "Genome sequence of Shigella flexneri 2a: insights into pathogenicity
RT through comparison with genomes of Escherichia coli K12 and O157.";
RN Nucleic Acids Res. 30:4432-4441(2002).
[5]
RN SEQUENCE FROM N.A.
RC SPECIES=S.flexneri; STRAIN=2457T / ATCC 700930 / Serotype 2a;
RX MEDLINE=22590274; PubMed=12704152;
RA Wei J., Goldberg M.B., Burland V., Venkatesan M.M., Deng W.,
RA Fournier G., Mayhew G.F., Plunkett G. III, Rose D.J., Darling A.,
RA Mau B., Perna N.T., Payne S.M., Runyen-Janecky L.J., Zhou S.,
RA Schwartz D.C., Blattner F.R.;
RT "Complete genome sequence and comparative genomics of Shigella
RT flexneri serotype 2a strain 2457T.";
RN Infect. Immun. 71:2775-2786(2003).
[6]
RN CHARACTERIZATION.
RC SPECIES=E.coli; STRAIN=K12;
RX MEDLINE=79082893; PubMed=215212;
RA Geis A., Plapp R.;
RT "Phospho-N-acetylmuramoyl-pentapeptide-transferase of Escherichia coli
RT K12. Properties of the membrane-bound and the extracted and partially
RT purified enzyme.";
RN Biochim. Biophys. Acta 527:414-424(1978).
[7]
RN SPECIES=E.coli; STRAIN=K12;
RC FUNCTION.
```

RA MEDLINE=91123172; PubMed=1846850;  
RA Ikeda M, Wachi M, Jung H.K., Ishino F., Matsushashi M.;  
RT "The Escherichia coli mray gene encoding UDP-N-acetylmuramoyl-  
pentapeptide: undecaprenyl-phosphate phospho-N-acetylmuramoyl-  
pentapeptide transferase.";  
RL J. Bacteriol. 173:1021-1026(1991).  
RN [8]  
RP TOPOLOGY.  
RC SPECIES=E. coli;  
RX MEDLINE=20022370; PubMed=10564498;  
RA Bouhass A., Mengin-Leclercq D., Le Beller D., Van Heijenoort J.;  
RT "Topological analysis of the mray protein catalyzing the first  
membrane step of peptidoglycan synthesis.";  
RL Mol. Microbiol. 34:576-585(1999).  
CC -1- FUNCTION: First step of the lipid cycle reactions in the  
CC biosynthesis of the cell wall peptidoglycan.  
CC -1- CATALYTIC ACTIVITY: UDPMur2Ac (Oyl-L-Ala-gamma-D-Ala-D-  
CC Glu-L-Lys-D-Ala-D-Ala)-diphosphoundecaprenol.  
CC Ala) + undecaprenyl phosphate = UMP + Mur2Ac(Oyl-L-Ala-gamma-D-  
CC Glu-L-Lys-D-Ala-D-Ala)-diphosphoundecaprenol.  
CC -1- COFACTOR: Magnesium or manganese.  
CC -1- PATHWAY: Peptidoglycan biosynthesis.  
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.  
CC -1- SIMILARITY: Belongs to the glycosyltransferase family 4. Mray  
CC subfamily.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; X51584; CAA35932.1; -;  
CC EMBL; X55034; CAA38864.1; -;  
CC EMBL; D10483; BAB96655.1; -;  
CC EMBL; AE000118; AAC73198.1; -;  
CC EMBL; AE015046; AAN41749.1; -;  
CC EMBL; AE016978; AAP15630.1; -;  
CC F.R. S08395; S08395.  
CC EcoGene; EG10604; mray.  
CC HAVAP; MF 00038; -; 1.  
CC InterPro; IPR00715; Glyco trans 4.  
CC InterPro; IPR003524; PNAcPdept trans.  
CC Pfam; PF00953; Glycos\_transf\_47.1.  
CC TIGRFAMs; TIGR00445; mray; 1.  
CC PROSITE; PS01347; MRAY\_1; 1.  
CC PROSITE; PS01348; MRAY\_2; 1.  
CC Peptidoglycan synthesis; Cell division; Transferrase; Transmembrane;  
CC Magnesium; Manganese; Complete proteome.  
CC DOMAIN 1 18 PERIPLASMIC.  
CC TRANSMEM 19 45  
CC DOMAIN 46 76 CYTOPLASMIC.  
CC TRANSMEM 77 90  
CC DOMAIN 91 96 PERIPLASMIC.  
CC TRANSMEM 97 113  
CC DOMAIN 114 133 CYTOPLASMIC.  
CC TRANSMEM 134 156  
CC DOMAIN 157 173 PERIPLASMIC.  
CC TRANSMEM 174 188  
CC DOMAIN 189 199 CYTOPLASMIC.  
CC TRANSMEM 200 220  
CC DOMAIN 221 238 PERIPLASMIC.  
CC TRANSMEM 239 251  
CC DOMAIN 252 270 CYTOPLASMIC.  
CC TRANSMEM 271 284  
CC DOMAIN 285 287 PERIPLASMIC.  
CC TRANSMEM 288 299  
CC DOMAIN 300 341  
CC TRANSMEM 342 357 CYTOPLASMIC.  
CC DOMAIN 358 360 PERIPLASMIC.  
CC SEQUENCE 360 AA; 39874 MW; F3550AFA3CD636AE CRC64;

Query Match 69.0%; Score 1268; DB 1; Length 360;  
Best Local Similarity 66.4%; Pred. No. 51e-83;  
Matches 239; Conservative 51; Mismatches 70; Indels 0; Gaps 0;  
QY 1 MLLLLASYLOQFYKGFVFOYLTIRGLLSVLTLSLSLWLGPMWIRTLOIPQIGQAVRND 60  
DB 1 MLVWLAELVKYISGFNFVSILTFRATVSLTLTALFTLSLWNGPRMIAHLQKLSFGQVVRND 60  
QY 61 GPQSHLSKKGTPMTGGALILTAITASTLLWADLSNRYVWVVLVVTLLFGAIGWYDDYRKV 120  
DB 61 GPESHFSKRGTPMTGGIMILTAIVISVLLWYFSNPVWCVLVVLVGYGVIGFVDDYRKV 120  
QY 121 IEKNSRGLPSRWKYFWOSVFGIGAAYELYNMTAETPIETTLIVPMLKSVETQLGIFVVL 180  
DB 121 VRKDTKGLIARWKYFWMVSIALGVAFALYLAGKDTPATQLVVPFFKDVMPQLGLFYLLA 180  
QY 181 YFVIVGSSNAVNLTDGLDGLAIMPTVMVAGALGFCVLSGNVKAFAEYLLIPNVPAGELI 240  
DB 181 YFVIVGSGNAVNLTDGLDGLAIMPTVFVAGGFALVAVATGNMFPASYLHIPYLRHAGELV 240  
QY 241 VFCAALVAGLGLFLEWNTYPAQVFMGDVGALALGAALGTIAVIVRQBIIVLIMGVVFVME 300  
DB 241 IVCTAIVGAGLGLFLEWNTYPAQVFMGDVGSLALGGALGIIAVLLRQEFLLVIMGVVFVE 300  
QY 301 TLSVMIQVASFKLTGRVFRMAPIHHPHFKGHPDPDRVIVRWTITVILVIGLATLKLK 360  
DB 301 TSLVILQVSGPLRGQIFRMAPIHHPHFKGHPDPDRVIVRWTITVILVIGLATLKLK 360  
Search completed: May 7, 2004, 08:38:11  
Job time : 19 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2004, 08:37:38 ; Search time 21 Seconds  
(without alignments)  
1648.998 Million cell updates/sec

Title: US-10-089-787-2  
Perfect score: 1837  
Sequence: 1 MLLLAELYQQFYKGFVQF.....RFWIITVLVIGLATLKLK 360

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 78:\*  
1: Pir1:\*  
2: Pir2:\*  
3: Pir3:\*  
4: Pir4:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1824	99.3	360	2 H83094	phospho-N-acetylmu
2	1296.5	70.6	361	2 H82763	phospho-N-acetylmu
3	1292	70.3	360	2 H82081	phospho-N-acetylmu
4	1268	69.0	360	1 S08295	phospho-N-acetylmu
5	1268	69.0	360	2 C90640	hypothetical prote
6	1268	69.0	360	2 C85491	hypothetical prote
7	1265	68.9	360	2 AE0068	phospho-N-acetylmu
8	1264	68.8	360	2 AB0518	phospho-N-acetylmu
9	1216	66.2	360	1 A64185	phospho-N-acetylmu
10	1100	59.9	360	2 E81200	phospho-N-acetylmu
11	1099	59.8	360	2 F81777	phospho-N-acetylmu
12	945	51.4	360	2 AB3324	phospho-N-acetylmu
13	917	49.9	366	2 AB2834	hypothetical prote
14	917	49.9	366	2 F97611	phospho-N-acetylmu
15	915.5	49.8	357	2 G84955	phospho-N-acetylmu
16	890.5	48.5	361	2 F97813	hypothetical prote
17	887.5	48.3	361	2 F71664	phospho-N-acetylmu
18	853.5	46.5	359	2 F70304	phospho-N-acetylmu
19	812.5	44.2	351	2 G70137	phospho-N-acetylmu
20	782.5	42.6	353	2 G81387	phospho-N-acetylmu
21	709.5	38.6	353	2 B71930	phospho-N-acetylmu
22	708.5	38.6	353	1 A64581	phospho-N-acetylmu
23	652	35.5	321	2 A89890	phospho-N-muramic
24	651	35.4	324	1 C47691	phospho-N-acetylmu
25	639	34.8	324	2 AE1329	phospho-N-acetylmu
26	636	34.6	324	2 AE1700	probable phospho-N
27	636	34.6	363	2 T34957	phospho-N-acetylmu
28	605.5	33.0	317	2 B97162	phospho-N-acetylmu
29	594.5	32.4	369	2 AE2345	phospho-N-acetylmu

RESULT 2

## RESULT 1

H83094  
phospho-N-acetylmuramoyl-pentapeptide- transferase PA4415 [imported] - Pseudomonas aeru:  
C:Species: Pseudomonas aeruginosa  
C:Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 31-Dec-2000  
C:Accession: H83094  
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Br:  
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lam,  
.; Lory, S.; Olson, M.V.  
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho:  
Nature 406, 959-964, 2000  
A:Reference number: A82950; MUID:20437337; PMID:10984043  
A:Accession: H83094  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-360 <STO>  
A:Cross-references: GB:AE004856; GB:AE004091; NID:G9950633; PIDN:AAG07803.1; GSPDB:GN001.  
A:Experimental source: strain PA01  
C:Genetics:  
A:Gene: mray; PA4415  
C:Superfamily: phospho-N-acetylmuramoyl-pentapeptide-transferase

Query Match 99.3%; Score 1824; DB 2; Length 360;  
Best Local Similarity 99.4%; Pred. No. 1.9e-130;  
Matches 358; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy	1	MLLLAEYLQOFYKGFVQFQYLTLRGILSVLTALSLSLWLGPMWIRTLOIQIGQAVRND	60
Db	1	MLLLAEYLQOFYKGFVQFQYLTLRGILSVLTALSLSLWLGPMWIRTLOIQIGQAVRND	60
Qy	61	GPQSHLSKKGTPMTGGALLITATAISTLLWADLSNRYVWVVLVVTLLFGAIGWDDYRKV	120
Db	61	GPQSHLSKKGTPMTGGALLITATAISTLLWADLSNRYVWVVLVVTLLFGAIGWDDYRKV	120
Qy	121	IEKNSRGLPSRWKYFQWSVFGIGAAVFLYMTAETPTETTLIVPMLKSVETQLGIFFWLT	180
Db	121	IEKNSRGLPSRWKYFQWSVFGIGAAVFLYMTAETPTETTLIVPMLKSVETQLGIFFWLT	180
Qy	181	YFVIVGSSNAVNLTGDLGLAIMPTVWAGALGIFCYLSGNVFAEYLLIPNVPFAGELI	240
Db	181	YFVIVGSSNAVNLTGDLGLAIMPTVWAGALGIFCYLSGNVFAEYLLIPNVPFAGELI	240
Qy	241	VFCALVAGLGLFWNTYPAQVFMGDVGALGAALGATIAIVROEIVLFIINGGVFVME	300
Db	241	VFCALVAGLGLFWNTYPAQVFMGDVGALGAALGATIAIVROEIVLFIINGGVFVME	300
Qy	301	TLVNIQVASFKLTRGRVFRMAPIHHHFELKGPDPVIVRPFWITVILVIGLATLKLK	360
Db	301	TLVNIQVASFKLTRGRVFRMAPIHHHFELKGPDPVIVRPFWITVILVIGLATLKLK	360

RESULT 2

B82763  
Phospho-N-acetylmuramoyl-pentapeptide- transferase XF0795 [imported] - Xylella fastidiosa  
C:Species: Xylella fastidiosa  
C>Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 02-Sep-2000  
C:Accession: B82763  
R:anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequencing  
Nature 406, 151-157, 2000  
A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.  
A:Reference number: A82515; MUID:20365717; PMID:10910347  
A:Note: for a complete list of authors see reference number A59328 below  
A:Accession: B82763  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-361 <STM>  
A:Cross-references: GB:A5003919; GB:A5003849; NID:9105675; PIDN:AAF83605.1; GSPDB:GN001  
A:Experimental source: strain 9a5C  
R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, H. as-Neto, E.; Docena, C.; El-Dorry, H.; Facincani, A.P.; Ferreira, A.J.S. submitted to GenBank, June 2000  
A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franco, S.C.; Franco, M.C.; Frohm J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laigh Chado, M.A.; Vadeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E.A.; Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.; F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A. Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveira M.; Tshukako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z A:Reference number: A59328  
A:Contents: annotation  
C:Genetics:  
A:Gene: XF0795  
C:Superfamily: phospho-N-acetylmuramoyl-pentapeptide-transferase  
Query Match 70.6%; Score 1296.5; DB 2; Length 361;  
Best Local Similarity 68.1%; Pred. No. 1.2e-90;  
Matches 246; Conservative 43; Mismatches 71; Indels 1; Gaps 1;  
Qy 1 MLLLLAEVLQPFYKGFVQVLTGRLSGLSVLTALSLSLWLGPMWIRTLQIPQIGQAVRND 60  
Db 1 MLFLARLWQFESLGLFNLTFRSLAALTALFLSLWTGIPVLIQKLSQFKGQPIQD 60  
Qy 61 GPQSHLSKGGTPTMGGAILTAIAISTLLMADLSNRYVWVLTLLFGAIGWDDYRKV 120  
Db 61 GPKHFSKAGTPTWGGSLILMTVLSLLNGDLNRYVWVLTLLFGAIGWDDYKIL 120  
Qy 121 IEKNSRGLPSRWKFWQSVFGIGAAVFLYMTAETPIETTLIVPMLKSVIEIQLGIFPVVL 179  
Db 121 ARDDPGLKRWKYLQSIQIFGLAAGLEFLYTDVPAAVTFYIPNFKSIALFLTSISFVAI 180  
Qy 180 TYFIVGSSNAVNLTDGLDLGAIIMPTVMVAGALGIFCYLSGNVKAFAEYLLIPNVPGAGEL 239  
Db 181 TYFIVGSSNAVNLTDGLDLGAIIMPTVMVAGALGIFCYLSGNVKAFAEYLLIPNVPGAGEL 240  
Qy 240 IVFCAALVAGLGLFWNTYPAQVFMGDVAGALGALGTIAIVRQEIIVLFIINGGVFVM 299  
Db 241 IIVCAIAGAGLGLFWNAYPAMVFMGDIGALGALGTIAIVRQELVVLVWVGGVFVI 300  
Qy 300 ETLSPVQVAFKLTGRVFRMAPIHHLKFWGPDPRVIVRFTITVLVILGLATLKL 359  
Db 301 ETLSPVQVAFKLTGRVFRMAPIHHLKFWGPDPRVIVRFTITVLVILGLATLKL 360  
Qy 360 R 360  
Db 361 R 361  
RESULT 3  
H82081  
phospho-N-acetylmuramoyl-pentapeptide- transferase VC2404 [imported] - Vibrio cholerae  
C:Species: Vibrio cholerae  
C>Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 02-Feb-2001  
C:Accession: H82081

R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.; Chardson, D.; Ermolaeva, M.D.; Vamathavan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, F. I.; R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.  
Nature 406, 477-483, 2000  
A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.  
A:Reference number: A82035; MUID:20406833; PMID:10952301  
A:Accession: H82081  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-360 <HEI>  
A:Cross-references: GB:A5004310; GB:A5003852; NID:9656963; PIDN:AAF95547.1; GSPDB:GN001  
A:Experimental source: serogroup O1; strain N16961; biotype El Tor  
C:Genetics:  
A:Gene: VC2404  
A:Map position: 1  
C:Superfamily: phospho-N-acetylmuramoyl-pentapeptide-transferase  
Query Match 70.3%; Score 1292; DB 2; Length 360;  
Best Local Similarity 66.9%; Pred. No. 2.7e-90;  
Matches 241; Conservative 51; Mismatches 68; Indels 0; Gaps 0;  
Qy 1 MLLLLAEVLQPFYKGFVQVLTGRLSGLSVLTALSLSLWLGPMWIRTLQIPQIGQAVRND 60  
Db 1 MIILAEVLQPFYKGFVQVLTGRLSGLSVLTALSLSLWLGPMWIRTLQIPQIGQAVRNE 60  
Qy 61 GPQSHLSKGGTPTMGGAILTAIAISTLLMADLSNRYVWVLTLLFGAIGWDDYRKV 120  
Db 61 GESHFSKRGTPMGVGMILAAITVLLMADLTPYVWVAVLLGAGVGFVDDYRKV 120  
Qy 121 IEKNSRGLPSRWKFWQSVFGIGAAVFLYMTAETPIETTLIVPMLKSVIEIQLGIFPVVL 180  
Db 121 VRKNTDGLIARWKYFWQSAIALVWAFALYAHQDTAAQTQVLPFKDVPGLGLMYILV 180  
Qy 181 YFIVGSSNAVNLTDGLDLGAIIMPTVMVAGALGIFCYLSGNVKAFAEYLLIPNVPGAGELI 240  
Db 181 YFIVGSSNAVNLTDGLDLGAIIMPTVMVAGALGIFCYLSGNVKAFAEYLLIPNVPGAGELI 240  
Qy 241 VFCALVAGLGLFWNTYPAQVFMGDVAGALGALGTIAIVRQEIIVLFIINGGVFVME 300  
Db 241 VVCTAMVAGLGLFWNTYPAQVFMGDVAGALGALGTIAIVRQEIIVLFIINGGVFVME 300  
Qy 301 TLSVMIQVAFKLTGRVFRMAPIHHLKFWGPDPRVIVRFTITVLVILGLATLKL 360  
Db 301 TLSVLIQVGSYKLRQIRFRMAPIHHLKFWGPDPRVIVRFTITVLVILGLATLKL 360  
RESULT 4  
S08395  
phospho-N-acetylmuramoyl-pentapeptide-transferase (EC 2.7.8.13) mray - Escherichia coli  
C:Species: Escherichia coli  
C>Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 01-Mar-2002  
C:Accession: S08395; S40597; G64730  
R:Ikeda, M.; Wachi, M.; Ishino, F.; Matsushashi, M.  
Nucleic Acids Res. 18, 1058, 1990  
A:Title: Nucleotide sequence involving mrd and an open reading frame ORF-Y spacing muf  
A:Reference number: S08395; MUID:90192099; PMID:2179861  
A:Accession: S08395  
A:Molecule type: DNA  
A:Residues: 1-360 <IKE>  
A:Cross-references: EMBL:X51584; NID:942058; PIDN:CAA35932.1; PID:942059  
R:Yura, T.; Mori, H.; Nagai, H.; Nagata, T.; Ishihama, A.; Fujita, N.; Isono, K.; Mizobu submitted to the EMBL Data Library, December 1992  
A:Description: Systematic sequencing of the Escherichia coli genome: analysis of the 0-2 Science 277, 1453-1462, 1997  
A:Reference number: S40531  
A:Accession: S40597  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-360 <YUR>  
A:Cross-references: EMBL:D10493; NID:9216434; PIDN:BA01352.1; PID:9216501  
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; C. A.; Rose, D.J.; Mau, B.; Shao, Y.  
Science 277, 1453-1462, 1997  
A:Title: The complete genome sequence of Escherichia coli K-12.



Db 241 IVCTAIVGAGLGLFWNTYPAQVFMGDSGALGALGIIAVLLRQEFFLLVINGGVFVVE 300  
Qy 301 TLSYMIQVASFGLTGRVFRMAPTHHFFELKSGPDPRIIVRFRWIIITVILVILGLATLKL 360  
Db 301 TLSVILQVSGFKLRGQRIFRMAPTHHFFELKSGPDPRIIVRFRWIIISLMLVILGLATLKL 360

RESULT 7  
AB0518  
phospho-N-acetylmuramoyl-pentapeptide-transferase (EC 2.7.8.13) [imported] - Yersinia pe  
C:Species: Yersinia pestis  
C>Date: 02-Nov-2001 #sequence\_revision 02-Nov-2001 #text\_change 27-Nov-2001  
C:Accession: AB0518  
R:Parkhill, J.; Kren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.  
geno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;  
il, M.; Kucherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrall,  
Nature 413, 523-527, 2001  
A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.  
A:Reference number: AB0001; MUID:21470413; PMID:11586360  
A:Accession: AB0068  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-360 <KUR>  
A:Cross-references: GB:AL590842; PIDN:CAC89408.1; PID:g15978644; GSPDB:GN00175  
C:Genetics:  
C:Superfamily: phospho-N-acetylmuramoyl-pentapeptide-transferase  
C:Keywords: transferase

Query Match 68.9%; Score 1265; DB 2; Length 360;  
Best Local Similarity 65.6%; Pred. No. 2.9e-88;  
Matches 236; Conservative 54; Mismatches 70; Indels 0; Gaps 0;

Qy 1 MLLLAELVLOQFYKGFQVLTGRLGILSVLTALSLSLWLPMPMIRTLQIPQIGQAVRND 60  
Db 1 MLVWLAELVLFKFSYSGNFVSLTFRALVSLTALFISLWGPRLARLQKLQIGQVVRND 60

Qy 61 GPQSHLSKGGPTMGALILTAISTLLMADLNRYVWVLTFLFAGLWDDIRKV 120  
Db 61 GPESHFSKRGPTMGGLMILFISLWGPRLARLQKLQIGQVVRND 60

Qy 121 LEKNSRGLPSRWKYPQVSGVFGIAGVFLYMTAETPIETLLVPMKSVLEIQLGIFFFVLT 180  
Db 121 VRKDTGLIARWKYFWSVIALGVAFALYVKGTPATQAVPFKDVMPQLGLFYLLS 180

Qy 181 YFVIYVSSNAVNLTDGLDGLAIMPTVMVAGALGIFCYLSGNVKAFAEYLLIPNVFAGELI 240  
Db 181 YFVIYVGTSGNAVNLTDGLDGLAIMPTVMVAGALGIFCYLSGNVKAFAEYLLIPNVFAGELI 240

Qy 241 VFCAALVAGLGLFWNTYPAQVFMGDSGALGALGIIAVLLRQEFFLLVINGGVFVVE 300  
Db 241 IVCTAIVGAGLGLFWNTYPAQVFMGDSGALGALGIIAVLLRQEFFLLVINGGVFVVE 300

Qy 301 TLSVMIQVASFGLTGRVFRMAPTHHFFELKSGPDPRIIVRFRWIIITVILVILGLATLKL 360  
Db 301 TLSVILQVSGFKLRGQRIFRMAPTHHFFELKSGPDPRIIVRFRWIIISLMLVILGLATLKL 360

RESULT 8  
AB0518  
phospho-N-acetylmuramoyl-pentapeptide-transferase (EC 2.7.8.13) - Haemophilus influenzae  
C:Species: Haemophilus influenzae  
C>Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999  
C:Accession: A64185  
R:Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, J.  
; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, C.  
; D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghegan, N.S.M.  
Science 269, 496-512, 1995  
A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter,  
A:Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.  
A:Reference number: A64000; MUID:95350630; PMID:7542800  
A:Accession: A64185  
A:Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-360 <TIGR>  
A:Cross-references: GB:U32793; GB:L42023; NID:g1574683; PIDN:AAC22790.1; PID:g1574690; 1  
C:Genetics:  
C:Gene: mray  
C:Function:  
C:Pathway: peptidoglycan biosynthesis  
C:Superfamily: phospho-N-acetylmuramoyl-pentapeptide-transferase  
C:Keywords: cell division; peptidoglycan biosynthesis; transferase; transmembrane protei  
P:29-45/Domain: transmembrane #status predicted <TM1>  
P:77-93/Domain: transmembrane #status predicted <TM2>  
P:98-114/Domain: transmembrane #status predicted <TM3>  
P:135-151/Domain: transmembrane #status predicted <TM4>  
P:172-188/Domain: transmembrane #status predicted <TM5>  
P:200-216/Domain: transmembrane #status predicted <TM6>  
P:239-255/Domain: transmembrane #status predicted <TM7>  
P:268-284/Domain: transmembrane #status predicted <TM8>  
P:289-305/Domain: transmembrane #status predicted <TM9>  
P:338-354/Domain: transmembrane #status predicted <TM10>

Db 241 IVCTAIVGAGLGLFWNTYPAQVFMGDSGALGALGIIAVLLRQEFFLLVINGGVFVVE 300  
Qy 301 TLSYMIQVASFGLTGRVFRMAPTHHFFELKSGPDPRIIVRFRWIIITVILVILGLATLKL 360  
Db 301 TLSVILQVSGFKLRGQRIFRMAPTHHFFELKSGPDPRIIVRFRWIIISLMLVILGLATLKL 360

RESULT 7  
AB0518  
phospho-N-acetylmuramoyl-pentapeptide-transferase (EC 2.7.8.13) [imported] - Yersinia pe  
C:Species: Yersinia pestis  
C>Date: 02-Nov-2001 #sequence\_revision 02-Nov-2001 #text\_change 27-Nov-2001  
C:Accession: AB0518  
R:Parkhill, J.; Kren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.  
geno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;  
il, M.; Kucherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrall,  
Nature 413, 523-527, 2001  
A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.  
A:Reference number: AB0001; MUID:21470413; PMID:11586360  
A:Accession: AB0068  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-360 <KUR>  
A:Cross-references: GB:AL590842; PIDN:CAC89408.1; PID:g15978644; GSPDB:GN00175  
C:Genetics:  
C:Superfamily: phospho-N-acetylmuramoyl-pentapeptide-transferase  
C:Keywords: transferase

Query Match 68.9%; Score 1265; DB 2; Length 360;  
Best Local Similarity 65.6%; Pred. No. 2.9e-88;  
Matches 236; Conservative 54; Mismatches 70; Indels 0; Gaps 0;

Qy 1 MLLLAELVLOQFYKGFQVLTGRLGILSVLTALSLSLWLPMPMIRTLQIPQIGQAVRND 60  
Db 1 MLVWLAELVLFKFSYSGNFVSLTFRALVSLTALFISLWGPRLARLQKLQIGQVVRND 60

Qy 61 GPQSHLSKGGPTMGALILTAISTLLMADLNRYVWVLTFLFAGLWDDIRKV 120  
Db 61 GPESHFSKRGPTMGGLMILFISLWGPRLARLQKLQIGQVVRND 60

Qy 121 LEKNSRGLPSRWKYPQVSGVFGIAGVFLYMTAETPIETLLVPMKSVLEIQLGIFFFVLT 180  
Db 121 VRKDTGLIARWKYFWSVIALGVAFALYVKGTPATQAVPFKDVMPQLGLFYLLS 180

Qy 181 YFVIYVSSNAVNLTDGLDGLAIMPTVMVAGALGIFCYLSGNVKAFAEYLLIPNVFAGELI 240  
Db 181 YFVIYVGTSGNAVNLTDGLDGLAIMPTVMVAGALGIFCYLSGNVKAFAEYLLIPNVFAGELI 240

Qy 241 VFCAALVAGLGLFWNTYPAQVFMGDSGALGALGIIAVLLRQEFFLLVINGGVFVVE 300  
Db 241 IVCTAIVGAGLGLFWNTYPAQVFMGDSGALGALGIIAVLLRQEFFLLVINGGVFVVE 300

Qy 301 TLSVMIQVASFGLTGRVFRMAPTHHFFELKSGPDPRIIVRFRWIIITVILVILGLATLKL 360  
Db 301 TLSVILQVSGFKLRGQRIFRMAPTHHFFELKSGPDPRIIVRFRWIIISLMLVILGLATLKL 360

RESULT 8  
AB0518  
phospho-N-acetylmuramoyl-pentapeptide-transferase (EC 2.7.8.13) - Haemophilus influenzae  
C:Species: Haemophilus influenzae  
C>Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999  
C:Accession: A64185  
R:Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, J.  
; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, C.  
; D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghegan, N.S.M.  
Science 269, 496-512, 1995  
A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter,  
A:Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.  
A:Reference number: A64000; MUID:95350630; PMID:7542800  
A:Accession: A64185  
A:Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-360 <TIGR>  
A:Cross-references: GB:U32793; GB:L42023; NID:g1574683; PIDN:AAC22790.1; PID:g1574690; 1  
C:Genetics:  
C:Gene: mray  
C:Function:  
C:Pathway: peptidoglycan biosynthesis  
C:Superfamily: phospho-N-acetylmuramoyl-pentapeptide-transferase  
C:Keywords: cell division; peptidoglycan biosynthesis; transferase; transmembrane protei  
P:29-45/Domain: transmembrane #status predicted <TM1>  
P:77-93/Domain: transmembrane #status predicted <TM2>  
P:98-114/Domain: transmembrane #status predicted <TM3>  
P:135-151/Domain: transmembrane #status predicted <TM4>  
P:172-188/Domain: transmembrane #status predicted <TM5>  
P:200-216/Domain: transmembrane #status predicted <TM6>  
P:239-255/Domain: transmembrane #status predicted <TM7>  
P:268-284/Domain: transmembrane #status predicted <TM8>  
P:289-305/Domain: transmembrane #status predicted <TM9>  
P:338-354/Domain: transmembrane #status predicted <TM10>

Query Match 66.2%; Score 1216; DB 1; Length 360;  
Best Local Similarity 61.9%; Pred. No. 1.5e-84;  
Matches 223; Conservative 63; Mismatches 74; Indels 0; Gaps 0;

QY 1 MLLLLAEYLQQFVKPGVFOYLTIRGILSVLTALSLSMLGPWMIRTLOIPOIGQAVRND 60  
DB :  
1 MLVLAELVLRYETAFAISYITVRANLALLTALFISLMWPKVIKRLQILKFGEVRND 60  
QY 61 GPQSHLSKKGTPTMGGALLTLTAISTLLWADLSNRYVVWLVTLLFGAIGWDDYRKV 120  
DB :  
61 GPESHFAKKGTPTMGGMWLFSGVSTLLWANLANPYIWCVLFVLFYGGAIGFDVDPRKI 120  
QY 121 IEKNSRGLPSRWKYFWQSVEFGCAAFVFLYMTAETPTETTLIVPMLKSVEIQLGIFPVLT 180  
DB :  
121 TRKNTDGLIARWKYFWMVSVALVAILMWLGHDTATLTVIPFFKDIMPQLGLEFIVLS 180  
QY 181 YFVIVGSSNAVNLTDLGLGIAMPTVMVAGALGIFCYLSGNVKFARYLLIPNVPGAGELI 240  
DB :  
181 YFVIVGTGNVNLTDGLGLIAMPTALVAGAFALIAWATGNVNFARYLHPIKYSESV 240  
QY 241 VFCAALVAGLAGLPENTYPAQVFMDGVGALAGAALGTIAIVROEIVLFTMGGVFVME 300  
DB :  
241 VFTALTVASGLFWNTYPAQVFMDGVGSALGGALGVVAIIVROEFFLVINGGVFVME 300  
QY 301 TLSVNIQVASFKLTRRRFRMAPIHHHFELKGNPDPRVIVREWIITVILVLIGLATLKLR 360  
DB :  
301 ALSVILQVGSKYKLRQIRFMAPIHHHFELKGNPEPRVIIREFWIIISMLVLMGLVTLKLR 360

RESULT 10  
EB1200  
phospho-N-acetylmuramoyl-pentapeptide- transferase NMB0418 [imported] - Neisseria mening  
C;Species: Neisseria meningitidis  
C;Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 19-Jan-2001  
C;Accession: EB1200  
R;Tettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A.  
Hikey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.;  
Ricci, H.; Qin, H.; Vanathevan, J.; Gill, J.; Scarlato, V.; Masignani, V.; Pizzi, M.  
Science 287, 1809-1815, 2000  
A;Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; Ve  
A;Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.  
A;Reference number: AB1000; UID:20175755; PMID:10710307  
A;Accession: EB1200  
A;Status: Preliminary  
A;Molecule type: DNA  
A;Residues: 1-360 <RET>  
A;Cross-references: GB:AE002098; GB:AE002398; GB:AE025640; PIDN:AAP40856.1; PID:g722564  
A;Experimental source: serogroup B, strain MC58  
C;Genetics:  
A;Gene: NMB0418  
C;Superfamily: phospho-N-acetylmuramoyl-pentapeptide-transferase

Query Match 59.9%; Score 1100; DB 2; Length 360;  
Best Local Similarity 58.2%; Pred. No. 8.4e-76;  
Matches 210; Conservative 59; Mismatches 90; Indels 2; Gaps 2;

QY 1 MLLLLABYLQOFYKGFVOYLTIRGILSVLTALSLSMLGPWMIRTLOIPOIGQAVRND 60  
DB 1 MFLLWAHF-SNWLTGLNFIQYTTTFRAVNAALTAALAFSLMFPTIRRLTKALKCGQAVRTD 59  
QY 61 GPQSHLSKKGTPTMGGALLTLTAISTLLWADLSNRYVVWLVTLLFGAIGWDDYRKV 120  
DB 60 GPQTHLVNGTPTMGSSLILTATVSTLLNGWNANPYIWLGLVLTATGALGFVDDRKY 119  
QY 121 IEKNSRGLPSRWKYFWQSVEFGICAAPVFLYMTAETPTETTLIVPMLKSVEIQLGIF-PVLT 179  
DB :  
120 VYKDPNGVSAKFKMVQSSVAITASLAFVLAANSANNLIIVPFKQIALPLGVGVFLV 179  
QY 180 TYFVIVGSSNAVNLTDLGLGIAMPTVMVAGALGFCVLSGNVKFARYLLIPNVPGAGEL 239  
DB :  
180 SYLTIVGTGNVNLTDGLDGUATFPVVLVAAGLAIFA YAGSHSQAOVLQLPYVAGANEV 239  
QY 240 IVFCAALVAGLAGLFWNTYPAQVFMDGVGALAGAALGTIAIVROEIVLFTMGGVFVME 299

**Db**

::    :      240 VIFCTAMCGACIGLFLWFNAPAQVFMGDVGALAGALTGVAVIVRQEFVLVINGGLFVV  300 ETLSVMIQVASFKLTGRVRFRMAPIHFFHPELKGDPPDRVIVRFWIITVILVLI GLATKL 309 EAVSVMLQGVIYKKTKKRI FLMAPIHHHYEQKGWKETQTVVVRFWIITVILVLI GLSTKI  360 R 360 +-----+ 360 R 360	<b>Dy</b>
--	-----------

**RESULT 11**  
phospho-N-acetylmuramoyl-pentapeptide-transferase [EC 2.7.8.13] NMA2066 [imported] - Neisseria meningitidis F81777  
C/Species: Neisseria meningitidis  
C/Date: 05-May-2000 #sequence\_revision 05-May-2000 #text\_change 02-Feb-2001  
C/Accession: F81777  
R/Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morell,  
S.; Holroyd, S.; Jagels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream,  
Nature 404, 502-506, 2000  
A>Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491.  
A/Reference number: AB1775; PMID:20222556; PMID:10761919  
A/Accession: F81777  
A>Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-360 <PAR>  
A/Cross-references: GB:AL162758; GB:AL157959; NID:G7380672; PIDN:CAB85284.1; PID:G7380694  
A/Experimental source: serogroup A, strain Z2491  
C/Genetics:  
A/Gene: mraX; NMA2066  
C/Superfamily: phospho-N-acetylmuramyl-pentapeptide-transferase  
C/Keywords: transferase

**Query Match**                  59.8%; Score 1099; DB 2; Length 360;  
**Best Local Similarity**      58.2%; Pred. No. le=75;  
**Matches**    210; Conservative    58; Mismatches    91; Indels         2; Gaps         2;

Dy 1 MLLLAELYQQFYKGFVGQYLTLRGILSVLTALSLSWLGPWMIRTLOIPQIGQAVRD 1 MFELAHFP-SNWLTQLNI FOYTTFRAVNAA LTALAFLMFGPWTIRRLTALKGGQA VRTD  61 GPQSLSKKGTPTWGGA LILTAISTLLWDLSNRKYVMVLVVTLLFCAIGHVD DRYKV 60 GPQTHLVKNHTPTWGS LLRTAI VTSTLLGNWANPYITILLGVLLATCALGYDDRWKV  121 IERNSGLPSSRKWFQWSV FGIGA AVFYMTAE TFIETT LIVPM LKSVE IQLGIF-FV VL 120 VKDPNGVS AKPKV WQSSVAIIAGLAFY LAANSANN ILIVPF FFKAIALPLGWGFLV L  180 TYFVIVGSSNA NVLTDL GDLAGIMPT VMVAVAGA LGFC YLSGNVKFAEYELLINVPV GA GE L 180 SYLTIVGTSNAVNL TDG LDG LGATFP PVLIVAGL AI FA VSASH SQPA QYL QL PPV AGAN EV  240 IVFCAA LVGAGL GF LFNTYP AQVF MG DV GALGA AL GTIA VI VIRQEI VLFIMG GV FVM 240 VIFCTAMCGAC IG LFLWN AP AQ VF MG DV GALGA AL GT VA IVRQE FVLV IMGGL FVV  300 ETLSVMIQVAS FKLT GR RV FR MAP IHFF HPELKGD PDPR VIV RF WI ITV ILV LI GL AT KL 309 EAVSV MLQGV IY KK TK KR IF LM API HH HY EQ KG WK ET Q TV VVR FW IITV ILV LI GL ST KI  360 R 360 +-----+ 360 R 360	<b>Dz</b>
--	-----------

**RESULT 12**  
AB3324  
phospho-N-acetylmuramoyl-pentapeptide-transferase [EC 2.7.8.13] [imported] - Brucella melitensis  
C/Species: Brucella melitensis  
C/Date: 01-Feb-2002 #sequence\_revision 01-Feb-2002 #text\_change 19-Apr-2002  
C/Accession: AB3324

R.; DelVecchio, V.G.; Kapatral, V.; Redkar, R.J.; Patra, G.; Mijer, C.; Los, T.; Ivanova, N.; Mazur, M.; Goldstein, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letes, Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002

A:Title: The genome sequence of the facultative intracellular pathogen *Brucella melitensis*

A:Reference number: AD3252; PMID:11756688

A:Accession: AB3324

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-360 <KUR>

A:Cross-references: GB:AE008917; PIDN:AAL51757.1; PID:g17982496; GSPDB:GN00190

A:Experimental source: strain 16M

C:Genetics:

A:Gene: BME10576

A:Map position: 1

C:Superfamily: phospho-N-acetylmuramoyl-pentapeptide-transferase

C:Keywords: transphospho

Query Match 51.4%; Score 945; DB 2; Length 360;  
Best Local Similarity 51.2%; Pred. No. 4.2e-64;  
Matches 185; Conservative 65; Mismatches 109; Indels 2; Gaps 2;

QY 1 MLLLAELYLQOFGVFOYLTIRGLSVLTALSLSLWLGPMWIRTLQIPQ-IGQAVRN 59  
DB 1 MMLFLTHFAEH-VTPENVFRYITRTGGAMITTSALIVFLFGPTIINSLRVKQKGPPIRA 59

QY 60 DGQSHLSKKGTPMGGLMILAGILGSSLLWGLSNVYVAVLVMTLGFCAIFYDDYK 119  
DB 60 DGPQTHFKAGTPTMGGLMILAGILGSSLLWGLSNVYVAVLVMTLGFCAIFYDDYK 119

QY 120 VIEKNRGLPSRWKYFWQSVFGL-----GAAVFLYM---TAETP-----IETTLIVPMLKS 167  
DB 120 VTKQSDKGFSGKAR-----LGIEFLTAATAVFFPMKQALASAPHGGLTSSIAFFPFKE 173

QY 168 VEIQLGIFFFVLTIVFVIGSSNANVLTGDLGLAIMPVTVAGALGIFCYLSGNVKFAEY 227  
DB 174 FVNLGFFVFLFGAFVIVGAGNAVLTGDLGLAIVPMIAAATFGVIAYLAGNAVAFNY 233

QY 228 LLIPNVFAGELIVFCAALVAGLGFVFNFTYPAQVFMGVDGALALGAALGTIAVIVRQE 287  
DB 234 LQINFVGTGELAVIVGAVIGAGLGFVFNFTYPAQVFMGVDGALALGAALGTIAVIVRQE 293

QY 288 IVLFIMGVFVMTLSVMIQVASFVLTGRRVFMPIHHFELKGPDPVIVRFWIITV 347  
DB 294 IVMVIVGGLFWFVMTLSVMIQVASFVLTGRRVFMPIHHFELKGPDPVIVRFWIITV 353

QY 348 ILVLIGLATLKL 360  
DB 354 GLALLGLATLKL 366

RESULT 14  
F97611

phospho-N-acetylmuramoyl-pentapeptide-transferase (UDP-murac-pentapeptide phosphotransferase)  
C:Species: Agrobacterium tumefaciens  
C:Date: 30-Sep-2001 #sequence\_revision 30-Sep-2001 #text\_change 18-Nov-2002  
C:Accession: F97611

R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman, A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.; Science 294, 2323-2328, 2001

A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent *Agrobacterium tumefaciens*

A:Reference number: A97359; PMID:11743194

A:Accession: F97611

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-366 <KUR>

A:Cross-references: GB:AE007869; PIDN:AAK87847.1; PID:g15157231; GSPDB:GN00169

C:Genetics:

A:Gene: AGR\_C\_3805

A:Map position: circular chromosome

C:Superfamily: phospho-N-acetylmuramoyl-pentapeptide-transferase

Query Match 49.9%; Score 917; DB 2; Length 366;  
Best Local Similarity 52.0%; Pred. No. 5.6e-62;  
Matches 194; Conservative 53; Mismatches 106; Indels 20; Gaps 6;

QY 1 MLLLAELYLQOFGVFOYLTIRGLSVLTALSLSLWLGPMWIRTLQIPQ-IGQAVRN 59  
DB 1 MMLFLTHFAEH-VTPENVFRYITRTGGAMITTSALIVFLFGPTIINSLRVKQKGPPIRA 59

QY 60 DGQSHLSKKGTPMGGLMILAGILGSSLLWGLSNVYVAVLVMTLGFCAIFYDDYK 119  
DB 60 DGPQTHFKAGTPTMGGLMILAGILGSSLLWGLSNVYVAVLVMTLGFCAIFYDDYK 119

QY 120 VIEKNRGLPSRWKYFWQSVFGL-----GAAVFLYM---TAETP-----IETTLIVPMLKS 167  
DB 120 VTKQSDKGFSGKAR-----LGIEFLTAATAVFFPMKQALASAPHGGLTSSIAFFPFKE 173

QY 168 VEIQLGIFFFVLTIVFVIGSSNANVLTGDLGLAIMPVTVAGALGIFCYLSGNVKFAEY 227

## RESULT 13

AB2834

hypothetical protein mray [imported] - Agrobacterium tumefaciens (strain C58, Dupont)

C:Species: Agrobacterium tumefaciens

C:Date: 11-Jan-2002 #sequence\_revision 11-Jan-2002 #text\_change 18-Nov-2002

C:Accession: AB2834

R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I.; Karp, P.; Romero, P.; Zhang, S. Science 294, 2317-2323, 2001

A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm, S.E.W.

A:Title: The Genome of the Natural Genetic Engineer *Agrobacterium tumefaciens* C58.

A:Reference number: AB2577; PMID:21608550; PMID:11743193

A:Accession: AB2834

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-366 <KUR>

A:Cross-references: GB:AE008688; PIDN:AAL43088.1; PID:g17740558; GSPDB:GN00186

A:Experimental source: strain C58 (Dupont)

C:Genetics:

A:Gene: mray

A:Map position: circular chromosome

Search completed: May 7, 2004, 08:40:55  
Job time : 22 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2004, 08:40:38 ; Search time 48 Seconds  
(without alignments)  
2081.751 Million cell updates/sec

Title: US-10-089-787-2

Perfect score: 1837

Sequence: 1 MLLLAELVQFQKFGVGFQ.....RFWITVLVLGLATLKLRL 360

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1140673 seqs, 277566755 residues

Total number of hits satisfying chosen parameters: 1140673

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA:\*

1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/2/pubpaa/US05\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/2/pubpaa/US09A\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*  
17: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1824	99.3	360	9	US-09-815-242-12011
2	1824	99.3	360	12	US-10-282-122A-66653
3	1699	91.9	360	12	US-10-282-122A-67901
4	1658	90.3	360	12	US-10-282-122A-69713
5	1292	70.3	360	12	US-10-282-122A-77459
6	1274	69.4	360	12	US-10-282-122A-66896
7	1268	69.0	360	9	US-09-741-669-451
8	1268	69.0	360	9	US-09-815-242-10028
9	1268	69.0	360	12	US-10-282-122A-56412
10	1265	68.9	360	12	US-10-282-122A-77920
11	1264	68.8	360	12	US-10-282-122A-76018
12	1261	68.6	360	12	US-10-282-122A-75186
13	1256	68.5	360	9	US-09-815-242-14075
14	1257	68.4	360	9	US-09-815-242-11691
15	1257	68.4	360	12	US-10-282-122A-55841

#### ALIGNMENTS

#### RESULT 1

US-09-815-242-12011  
Sequence 12011, Application US/09815242  
; Patent No. US20020061569A1  
; GENERAL INFORMATION:  
; APPLICANT: Hasebeck, Robert  
; APPLICANT: Ohlsen, Kari L.  
; APPLICANT: Zyskind, Judith W.  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John D.  
; APPLICANT: Carr, Grant J.  
; APPLICANT: Yamamoto, Robert T.  
; APPLICANT: Xu, H. Howard  
; TITLE OF INVENTION: Identification of Essential Genes in  
; FILE REFERENCE: ELITRA.011A  
; CURRENT APPLICATION NUMBER: US/09/815,242  
; CURRENT FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: 60/191,078  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: 60/206,848  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR APPLICATION NUMBER: 60/207,727  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: 60/242,578  
; PRIOR FILING DATE: 2000-10-23  
; PRIOR APPLICATION NUMBER: 60/253,625  
; PRIOR FILING DATE: 2000-11-27  
; PRIOR APPLICATION NUMBER: 60/257,931  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: 60/269,308  
; PRIOR FILING DATE: 2001-02-16  
; NUMBER OF SEQ ID NOS: 14110  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 12011  
; LENGTH: 360  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa

16 1257 68.4 360 12 US-10-282-122A-59586  
17 1250 68.0 360 12 US-10-282-122A-68814  
18 1228 66.8 352 12 US-10-282-122A-73233  
19 1216 66.2 360 9 US-09-815-242-11186  
20 1216 66.2 360 12 US-10-282-122A-58419  
21 1205.5 65.6 389 12 US-10-282-122A-50013  
22 1202.5 65.5 389 12 US-10-282-122A-51377  
23 1185.5 64.5 389 12 US-10-282-122A-47105  
24 1118 60.9 372 12 US-10-282-122A-44716  
25 1099 59.8 360 12 US-10-282-122A-66064  
26 1094 59.6 360 12 US-10-282-122A-65130  
27 972.5 52.9 365 12 US-10-282-122A-63058  
28 901.5 49.1 339 12 US-10-282-122A-61467  
29 812.5 44.2 351 12 US-10-282-122A-47085  
30 782.5 42.6 333 12 US-10-282-122A-54319  
31 755.5 41.1 215 12 US-10-282-122A-48088  
32 709.5 38.6 353 12 US-10-335-977-5598  
33 709.5 38.6 376 12 US-10-335-977-5598  
34 708.5 38.6 353 9 US-09-815-242-11328  
35 708.5 38.6 353 12 US-10-282-122A-58718  
36 671 36.5 324 12 US-10-282-122A-46783  
37 652 35.5 321 9 US-09-815-242-12347  
38 652 35.5 321 9 US-09-815-242-12794  
39 652 35.5 321 12 US-10-282-122A-44257  
40 642 34.9 317 12 US-10-282-122A-52288  
41 639 34.8 324 12 US-10-282-122A-60973  
42 637 34.7 422 12 US-10-282-122A-48543  
43 636 34.6 321 12 US-10-282-122A-70722  
44 629.5 34.3 321 12 US-10-282-122A-71642  
45 617.5 33.6 357 14 US-10-156-761-13646

Sequence 59586, A  
Sequence 68814, A  
Sequence 73233, A  
Sequence 11186, A  
Sequence 58419, A  
Sequence 50013, A  
Sequence 51377, A  
Sequence 49105, A  
Sequence 44716, A  
Sequence 66064, A  
Sequence 65130, A  
Sequence 63058, A  
Sequence 61467, A  
Sequence 47085, A  
Sequence 54319, A  
Sequence 48088, A  
Sequence 5598, Ap  
Sequence 11328, A  
Sequence 58718, A  
Sequence 46783, A  
Sequence 12347, A  
Sequence 12794, A  
Sequence 44257, A  
Sequence 52288, A  
Sequence 60973, A  
Sequence 48543, A  
Sequence 70722, A  
Sequence 71642, A  
Sequence 13646, A

US-09-815-242-12011

Query Match 99.3%; Score 1824; DB 9; Length 360;  
Best Local Similarity 99.4%; Pred. No. 6.9e-166;  
Matches 358; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
QY 1 MLLLAAYLQOYKFGVFOYLTGRLSGLTSLTSLSLMLGPMWIRTLQIQTQAVRND 60  
DB 1 MLLLAAYLQOYKFGVFOYLTGRLSGLTSLTSLSLMLGPMWIRTLQIQTQAVRND 60  
QY 61 GPQSHLSKGTPTMGGALILTAISTLLWADLSNRVYVVLVTLFAGIHWDDYRKV 120  
DB 61 GPQSHLSKGTPTMGGALILTAISTLLWADLSNRVYVVLVTLFAGIHWDDYRKV 120  
QY 121 IEKNSRGLPSRWKFWQSVFGIGAAFLYMTAETPIETTLIVPMLKSVEIQIGFFVLT 180  
DB 121 IEKNSRGLPSRWKFWQSVFGIGAAFLYMTAETPIETTLIVPMLKSVEIQIGFFVLT 180  
QY 181 YFVIVGSSNAVNLTDGLDLAIMPTVMVAGALGIFCYLSGNVFAEYLLIPNVPGAGELI 240  
DB 181 YFVIVGSSNAVNLTDGLDLAIMPTVMVAGALGIFCYLSGNVFAEYLLIPNVPGAGELI 240  
QY 241 VFCAALVGAGLGFLEWNTYPAQVFMGVDGALALGAALGTTIAVIRQEIIVLFGVGVFVME 300  
DB 241 VFCAALVGAGLGFLEWNTYPAQVFMGVDGALALGAALGTTIAVIRQEIIVLFGVGVFVME 300  
QY 301 TLSVMIQVASFKLTCRRVFRMAPIHHPHFELKGPDPVIVRFFWITVILVIGLATLKL 360  
DB 301 TLSVMIQVASFKLTCRRVFRMAPIHHPHFELKGPDPVIVRFFWITVILVIGLATLKL 360

RESULT 2

US-10-282-122A-66653  
; Sequence 66653, Application US/10282122A  
; Publication No. US20040029129A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Liangsu  
; APPLICANT: Zamudio, Carlos  
; APPLICANT: Malone, Cheryl  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Kari  
; APPLICANT: Zyskind, Judith  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John  
; APPLICANT: Carr, Grant  
; APPLICANT: Yamamoto, Robert  
; APPLICANT: Forsyth, R.  
; APPLICANT: Xu, H.  
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
; FILE REFERENCE: ELITRA.034A  
; CURRENT FILING DATE: 2003-02-20  
; PRIOR FILING DATE: 2003-02-20  
; PRIOR APPLICATION NUMBER: 60/191,078  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: 60/206,848  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR APPLICATION NUMBER: 60/207,727  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR APPLICATION NUMBER: 60/230,335  
; PRIOR FILING DATE: 2000-09-06  
; PRIOR APPLICATION NUMBER: 60/230,347  
; PRIOR FILING DATE: 2000-09-06  
; PRIOR APPLICATION NUMBER: 60/230,335  
; PRIOR FILING DATE: 2000-09-09  
; PRIOR APPLICATION NUMBER: 60/207,727  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: 60/230,335  
; PRIOR FILING DATE: 2000-09-06  
; PRIOR APPLICATION NUMBER: 60/230,347  
; PRIOR FILING DATE: 2000-09-09  
; PRIOR APPLICATION NUMBER: 60/242,578  
; PRIOR FILING DATE: 2000-10-23  
; PRIOR APPLICATION NUMBER: 60/253,625  
; PRIOR FILING DATE: 2000-11-27  
; PRIOR APPLICATION NUMBER: 60/257,931  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: 60/267,636  
; PRIOR FILING DATE: 2001-02-09  
; PRIOR APPLICATION NUMBER: 60/269,308  
; PRIOR FILING DATE: 2001-02-16  
; Remaining Prior Application data removed - See File Wrapper or PALM.

NUMBER OF SEQ ID NOS: 78614  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 66653  
; LENGTH: 360  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-10-282-122A-66653

Query Match 99.3%; Score 1824; DB 12; Length 360;  
Best Local Similarity 99.4%; Pred. No. 6.9e-166;  
Matches 358; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
QY 1 MLLLAAYLQOYKFGVFOYLTGRLSGLTSLTSLSLMLGPMWIRTLQIQTQAVRND 60  
DB 1 MLLLAAYLQOYKFGVFOYLTGRLSGLTSLTSLSLMLGPMWIRTLQIQTQAVRND 60  
QY 61 GPQSHLSKGTPTMGGALILTAISTLLWADLSNRVYVVLVTLFAGIHWDDYRKV 120  
DB 61 GPQSHLSKGTPTMGGALILTAISTLLWADLSNRVYVVLVTLFAGIHWDDYRKV 120  
QY 121 IEKNSRGLPSRWKFWQSVFGIGAAFLYMTAETPIETTLIVPMLKSVEIQIGFFVLT 180  
DB 121 IEKNSRGLPSRWKFWQSVFGIGAAFLYMTAETPIETTLIVPMLKSVEIQIGFFVLT 180  
QY 181 YFVIVGSSNAVNLTDGLDLAIMPTVMVAGALGIFCYLSGNVFAEYLLIPNVPGAGELI 240  
DB 181 YFVIVGSSNAVNLTDGLDLAIMPTVMVAGALGIFCYLSGNVFAEYLLIPNVPGAGELI 240  
QY 241 VFCAALVGAGLGFLEWNTYPAQVFMGVDGALALGAALGTTIAVIRQEIIVLFGVGVFVME 300  
DB 241 VFCAALVGAGLGFLEWNTYPAQVFMGVDGALALGAALGTTIAVIRQEIIVLFGVGVFVME 300  
QY 301 TLSVMIQVASFKLTCRRVFRMAPIHHPHFELKGPDPVIVRFFWITVILVIGLATLKL 360  
DB 301 TLSVMIQVASFKLTCRRVFRMAPIHHPHFELKGPDPVIVRFFWITVILVIGLATLKL 360

RESULT 3

US-10-282-122A-67901  
; Sequence 67901, Application US/10282122A  
; Publication No. US20040029129A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Liangsu  
; APPLICANT: Zamudio, Carlos  
; APPLICANT: Malone, Cheryl  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Kari  
; APPLICANT: Zyskind, Judith  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John  
; APPLICANT: Carr, Grant  
; APPLICANT: Yamamoto, Robert  
; APPLICANT: Forsyth, R.  
; APPLICANT: Xu, H.  
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
; FILE REFERENCE: ELITRA.034A  
; CURRENT FILING DATE: 2003-02-20  
; PRIOR FILING DATE: 2003-02-20  
; PRIOR APPLICATION NUMBER: 60/191,078  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: 60/206,848  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR APPLICATION NUMBER: 60/207,727  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: 60/230,335  
; PRIOR FILING DATE: 2000-09-06  
; PRIOR APPLICATION NUMBER: 60/230,347  
; PRIOR FILING DATE: 2000-09-09  
; PRIOR APPLICATION NUMBER: 60/242,578  
; PRIOR FILING DATE: 2000-10-23  
; PRIOR APPLICATION NUMBER: 60/253,625  
; PRIOR FILING DATE: 2000-11-27  
; PRIOR APPLICATION NUMBER: 60/257,931

;; PRIOR FILING DATE: 2000-12-22  
;; PRIOR APPLICATION NUMBER: 60/267,636  
;; PRIOR FILING DATE: 2001-02-09  
;; PRIOR APPLICATION NUMBER: 60/269,308  
;; PRIOR FILING DATE: 2001-02-16  
;; Remaining Prior Application data removed - See File Wrapper or PALM.  
;; NUMBER OF SEQ ID NOS: 78614  
;; SOFTWARE: PatentIn version 3.1  
;; SEQ ID NO 67901  
;; LENGTH: 360  
;; TYPE: PRT  
;; ORGANISM: Pseudomonas putida  
US-10-282-122A-67901

Query Match 91.9%; Score 1689; DB 12; Length 360;  
Best Local Similarity 90.3%; Pred. No. 5.6e-153;  
Matches 325; Conservative 18; Mismatches 17; Indels 0; Gaps 0;  
QY 1 MLLLAELVQLQYKGFQYLTGILSVLTALSLSLWLGPMWIRTIQIQIGAVRND 60  
DB 1 MLLLAELVQLQYKGFQYLTGILSVLTALSLSLWLGPMWIRTIQIQIGAVRND 60  
QY 61 GPQSHLSKGTPTMGALILTAISTLLWADLSNRYVWVLTLLFCAIGWDDYRKV 120  
DB 61 GPQSHLSKGTPTMGALILTAISTLLWADLSNRYVWVLTLLFCAIGWDDYRKV 120  
QY 121 IEKNSRGLPSRMKYFQWSVFGIGAAVFLYMTAETPIETTLIVPMLKSVETIOLGIFVVL 180  
DB 121 IEKNSRGLPSRMKYFQWSVFGIGAAVFLYMTAETPIETTLIVPMLKSVETIOLGIFVVL 180  
QY 181 YFVIVGSSNAVNLTDGLGLAIMPVTWVAGALGIFCYLSGNVKFAYLLIPNVPAGAGELI 240  
DB 181 YFVIVGSSNAVNLTDGLGLAIMPVTWVAGALGIFCYLSGNVKFAYLLIPNVPAGAGELI 240  
QY 241 VFCALVAGLGLFLENFTYPAQVFMGDVGALGALGTIAVIVROEIVLFIIMGVFMVE 300  
DB 241 VFCALVAGLGLFLENFTYPAQVFMGDVGALGALGTIAVIVROEIVLFIIMGVFMVE 300  
QY 301 TLSVMIQVASFKLTKRRVFRMAPIHFFELKGPDPRIIVRFWIIITVILVILGLATLKL 360  
DB 301 TLSVMIQVASFKLTKRRVFRMAPIHFFELKGPDPRIIVRFWIIITVILVILGLATLKL 360

RESULT 4  
US-10-282-122A-69713  
; Sequence 69713, Application US/10282122A  
; Publication No. US20040029129A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Liangsu  
; APPLICANT: Zamudio, Carlos  
; APPLICANT: Malone, Cheryl  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Kari  
; APPLICANT: Zyskind, Judith  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John  
; APPLICANT: Carr, Grant  
; APPLICANT: Yamamoto, Robert  
; APPLICANT: Forsyth, R.  
; APPLICANT: Xu, H.  
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
; FILE REFERENCE: ELITRA.034A  
; CURRENT APPLICATION NUMBER: US/10/282,122A  
; CURRENT FILING DATE: 2003-02-20  
; PRIOR APPLICATION NUMBER: 60/191,078  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: 60/206,848  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR APPLICATION NUMBER: 60/207,727  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: 60/230,335  
; PRIOR FILING DATE: 2000-09-06  
; PRIOR APPLICATION NUMBER: 60/230,347

;; PRIOR FILING DATE: 2000-09-09  
;; PRIOR APPLICATION NUMBER: 60/242,578  
;; PRIOR FILING DATE: 2000-10-23  
;; PRIOR APPLICATION NUMBER: 60/253,625  
;; PRIOR FILING DATE: 2000-11-27  
;; PRIOR APPLICATION NUMBER: 60/257,931  
;; PRIOR FILING DATE: 2000-12-22  
;; PRIOR APPLICATION NUMBER: 60/267,636  
;; PRIOR FILING DATE: 2001-02-09  
;; PRIOR APPLICATION NUMBER: 60/269,308  
;; Remaining Prior Application data removed - See File Wrapper or PALM.  
;; NUMBER OF SEQ ID NOS: 78614  
;; SOFTWARE: PatentIn version 3.1  
;; SEQ ID NO 69713  
;; LENGTH: 360  
;; TYPE: PRT  
;; ORGANISM: Pseudomonas syringae  
US-10-282-122A-69713

Query Match 90.3%; Score 1658; DB 12; Length 360;  
Best Local Similarity 89.7%; Pred. No. 5.1e-150;  
Matches 323; Conservative 17; Mismatches 20; Indels 0; Gaps 0;  
QY 1 MLLLAELVQLQYKGFQYLTGILSVLTALSLSLWLGPMWIRTIQIQIGAVRND 60  
DB 1 MLLLAELVQLQYKGFQYLTGILSVLTALSLSLWLGPMWIRTIQIQIGAVRND 60  
QY 61 GPQSHLSKGTPTMGALILTAISTLLWADLSNRYVWVLTLLFCAIGWDDYRKV 120  
DB 61 GPQSHLSKGTPTMGALILTAISTLLWADLSNRYVWVLTLLFCAIGWDDYRKV 120  
QY 121 IEKNSRGLPSRMKYFQWSVFGIGAAVFLYMTAETPIETTLIVPMLKSVETIOLGIFVVL 180  
DB 121 IEKNSRGLPSRMKYFQWSVFGIGAAVFLYMTAETPIETTLIVPMLKSVETIOLGIFVVL 180  
QY 181 YFVIVGSSNAVNLTDGLGLAIMPVTWVAGALGIFCYLSGNVKFAYLLIPNVPAGAGELI 240  
DB 181 YFVIVGSSNAVNLTDGLGLAIMPVTWVAGALGIFCYLSGNVKFAYLLIPNVPAGAGELI 240  
QY 241 VFCALVAGLGLFLENFTYPAQVFMGDVGALGALGTIAVIVROEIVLFIIMGVFMVE 300  
DB 241 VFCALVAGLGLFLENFTYPAQVFMGDVGALGALGTIAVIVROEIVLFIIMGVFMVE 300  
QY 301 TLSVMIQVASFKLTKRRVFRMAPIHFFELKGPDPRIIVRFWIIITVILVILGLATLKL 360  
DB 301 TLSVMIQVASFKLTKRRVFRMAPIHFFELKGPDPRIIVRFWIIITVILVILGLATLKL 360

RESULT 5  
US-10-282-122A-77459  
; Sequence 77459, Application US/10282122A  
; Publication No. US20040029129A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Liangsu  
; APPLICANT: Zamudio, Carlos  
; APPLICANT: Malone, Cheryl  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Kari  
; APPLICANT: Zyskind, Judith  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John  
; APPLICANT: Carr, Grant  
; APPLICANT: Yamamoto, Robert  
; APPLICANT: Forsyth, R.  
; APPLICANT: Xu, H.  
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
; FILE REFERENCE: ELITRA.034A  
; CURRENT APPLICATION NUMBER: US/10/282,122A  
; CURRENT FILING DATE: 2003-02-20  
; PRIOR APPLICATION NUMBER: 60/191,078  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: 60/206,848

```
/ PRIOR FILING DATE: 2000-05-23
/ PRIOR APPLICATION NUMBER: 60/207,727
/ PRIOR FILING DATE: 2000-05-26
/ PRIOR APPLICATION NUMBER: 60/230,335
/ PRIOR FILING DATE: 2000-09-06
/ PRIOR APPLICATION NUMBER: 60/230,347
/ PRIOR FILING DATE: 2000-09-09
/ PRIOR APPLICATION NUMBER: 60/242,578
/ PRIOR FILING DATE: 2000-10-23
/ PRIOR APPLICATION NUMBER: 60/253,625
/ PRIOR FILING DATE: 2000-11-27
/ PRIOR APPLICATION NUMBER: 60/257,931
/ PRIOR FILING DATE: 2000-12-22
/ PRIOR APPLICATION NUMBER: 60/267,636
/ PRIOR FILING DATE: 2001-02-09
/ PRIOR APPLICATION NUMBER: 60/269,308
/ PRIOR FILING DATE: 2001-02-16
/ Remaining Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 78614
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 77459
/ LENGTH: 360
/ TYPE: PRT
/ ORGANISM: Vibrio cholerae
/ US-10-282-122A-77459

Query Match          70.3%; Score 1292; DB 12; Length 360;
Best Local Similarity 66.9%; Pred. No. 5.2e-115;
Matches 241; Conservative 51; Mismatches 68; Indels 0; Gaps 0;

Qy 1 MLLLLAEYLQOYFGVGFQYLTGILSVLTALSLSLWLGPMWIRTLQIPQIGQAVRND 60
Db 1 MIWLAELLOPYFFSFFLFEYLSFRAIVSILTALGSLWNGPRMIRKQLQWLGQVVRNE 60
Qy 61 GPQSHLSKGGTPTMGGAIIILTAISTLLWADLSNRVWVVLVLTGALGHWDDYRKV 120
Db 61 GPESHFSKRGTPMGVWMLAIIITVLLWADLTNPVMAVLAVLGYGAVGVDDYRKV 120
Qy 121 IEKNSRGLPSWKYFWQSVFGIGAAFLYMTAETPIETTLIVPMLKSVETQLGIFVLT 180
Db 121 VRKNTDGLIARWKYFWQSAIYVAFALYAHQDQTAATQLVWFFKDWMPQLGLMIVLT 180
Qy 181 YFVIVGSSNAVNLTDGLDLAIMPTVMVAGALGIFCYLSGNVKAFAEYLLIPNVPAGELI 240
Db 181 YFVIVGTSNAVNLTDGLDLAIMPTVLVAGAFALIAWATGNVNFAYLHPIYHSELV 240
Qy 241 VFCAALVAGLGLFWNTYPAQVFMGVDGALGALGTTIAIVRQBIIVFIMGVVFVME 300
Db 241 VVCTAMVAGLGLFWNTYPAQVFMGVDGALGALGTTIAIVRQBFVLVIMGVVFVME 300
Qy 301 TLSVMIQVASFKLGRVFRMAPIHHPHFKGWDPDRVIVRFTIITVLVLIGLATLKL 360
Db 301 TLSVILQVGSYKLRQIRFRMAPIHHPHFKGWPEPRVIVRFTIISVLVLIGLATLKL 360

RESULT 6
US-10-282-122A-66896
/ Sequence 66896, Application US/10282122A
/ Publication No. US20040029129A1
/ GENERAL INFORMATION:
/ APPLICANT: Wang, Liangsu
/ APPLICANT: Zamudio, Carlos
/ APPLICANT: Malone, Cheryl
/ APPLICANT: Haselbeck, Robert
/ APPLICANT: Olsen, Kari
/ APPLICANT: Zyskind, Judith
/ APPLICANT: Wall, Daniel
/ APPLICANT: Trawick, John
/ APPLICANT: Carr, Grant
/ APPLICANT: Yamamoto, Robert
/ APPLICANT: Forsyth, R.
/ APPLICANT: Xu, H.
/ TITLE OF INVENTION: Identification of Essential Genes in Microorganisms

/ PRIOR FILING DATE: 2000-05-23
/ PRIOR APPLICATION NUMBER: 60/207,727
/ PRIOR FILING DATE: 2000-05-26
/ PRIOR APPLICATION NUMBER: 60/230,335
/ PRIOR FILING DATE: 2000-09-06
/ PRIOR APPLICATION NUMBER: 60/230,347
/ PRIOR FILING DATE: 2000-09-09
/ PRIOR APPLICATION NUMBER: 60/242,578
/ PRIOR FILING DATE: 2000-10-23
/ PRIOR APPLICATION NUMBER: 60/253,625
/ PRIOR FILING DATE: 2000-11-27
/ PRIOR APPLICATION NUMBER: 60/257,931
/ PRIOR FILING DATE: 2000-12-22
/ PRIOR APPLICATION NUMBER: 60/267,636
/ PRIOR FILING DATE: 2001-02-09
/ PRIOR APPLICATION NUMBER: 60/269,308
/ PRIOR FILING DATE: 2001-02-16
/ Remaining Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 78614
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 77459
/ LENGTH: 360
/ TYPE: PRT
/ ORGANISM: Pasteurella multocida
/ US-10-282-122A-66896

Query Match          69.4%; Score 1274; DB 12; Length 360;
Best Local Similarity 65.8%; Pred. No. 2.7e-113;
Matches 237; Conservative 57; Mismatches 66; Indels 0; Gaps 0;

Qy 1 MLLLLAEYLQOYFGVGFQYLTGILSVLTALSLSLWLGPMWIRTLQIPQIGQAVRND 60
Db 1 MLVWIGEFLQOYYSFNFVISITVRAILLALTALLSVLWIGPKVIRRLQLKFGQGVHRD 60
Qy 61 GPQSHLSKGGTPTMGGAIIILTAISTLLWADLSNRVWVVLVLTGALGHWDDYRKV 120
Db 61 GPESHFSKRGTPMGVWMLFAITVSTLLWANLANPYWVFSLVLLGYGALGFVDDYRKI 120
Qy 121 IEKNSRGLPSWKYFWQSVFGIGAAFLYMTAETPIETTLIVPMLKSVETQLGIFVLT 180
Db 121 TRKNTDGLIARWKYFWLSVIALVAAGFYATGKDTATRLVVPFFKEMPQLGFIILT 180
Qy 181 YFVIVGSSNAVNLTDGLDLAIMPTVMVAGALGIFCYLSGNVKAFAEYLLIPNVPAGELI 240
Db 181 YFVIVGTSNAVNLTDGLDLAIMPTVLVAGAFALIAWATGNVNFAYLHPIYKFSBELV 240
Qy 241 VFCAALVAGLGLFWNTYPAQVFMGVDGALGALGTTIAIVRQBIIVFIMGVVFVME 300
Db 241 VFCTAIVAGLGLFWNTYPAQVFMGVDGSLGALGALGVAIVRQBEFLVIMGVVFVME 300
Qy 301 TLSVMIQVASFKLGRVFRMAPIHHPHFKGWDPDRVIVRFTIITVLVLIGLATLKL 360
Db 301 TLSVILQVGSYKLRQIRFRMAPIHHPHFKGWPEPRVIVRFTIISMLVLIGLVLTKL 360

RESULT 7
US-09-741-669-451
/ Sequence 451, Application US/09741669
/ Patent No. US2002022718A1
/ GENERAL INFORMATION:
/ APPLICANT: Forsyth, R. Allyn
/ APPLICANT: Ohlsen, Kari L.
/ APPLICANT: Zyskind, Judith W.
/ TITLE OF INVENTION: Genes identified as required for
/ proliferation of E. coli
/ FILE REFERENCE: ELITRA.009A
/ CURRENT APPLICATION NUMBER: US/09/741,669
```

```

; CURRENT FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: US 60/173005
; PRIOR FILING DATE: 1999-12-23
; NUMBER OF SEQ ID NOS: 481
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 451
; LENGTH: 360
; TYPE: PRY
; ORGANISM: Escherichia coli
; US-09-741-669-451

Query Match          69.0%; Score 1268; DB 9; Length 360;
Best Local Similarity 66.4%; Pred. No.le=112;
Matches 239; Conservative 51; Mismatches 70; Indels 0; Gaps 0

Qy      1 MLLLAEYLQOIFYKGFGVQLTURGILSVLTALSUSLNLGPMWIRTLPQIPQAVRND 60
        |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::
Db      1 MLVLVAELHVKYISGFNVFSYTLFRAIVSLTALFISLMWGPRIAHLOKLSFGQVRND 60

Qy      61 GPQSHLSKKGTPTMGGALLITAIISTLLWADISNRVVVVLVVTLFCAGIWNDDYRKV 120
        |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::
Db      61 GESHFSGRGTPPTMGGINMLTAIVISLLWAYSPNPYVCVLVLVGYGIVGVDDYRKV 120

Qy      121 IENSRGLPSRWKYFMQVFGIGAAYFLYMTAETPIETTLLIVPMLKSVEIQIGIFFVVLVT 180
        |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::
Db      121 VRKDTKGLIARWKYFMSVSIALGVAPALYLAGXDTPATQLVVPFFSDVMPLGLGVILLA 180

Qy      181 YFVIYVGSNANVLLTDGLDGLAIMTVWVAGALGIFCYLSGNWKFAFYILLINPVPGAGELI 240
        |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::
Db      181 YFVIYVGTGNVNLNTDGLDGLAINPTVFVAGFALVAWATGMNNMFASLIHPYLRHAGSLV 240

Qy      241 VFCAALVAGALGFLMFTNYPAQVPMGDVGALGAALGTTAVTVROEIYLVLFIMGGVYME 300
        |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::
Db      241 IVCTAIVANGLGFLMFTNYPAQVPMGDVGSALGGALGITAVLLRDEFFLIVIMGGVFVE 300

Qy      301 TTSVMNIQVASFKLTGRRVFRMAPITHHFELKGMPPDRVIVRFWIIITVLVLIGLATKLRL 360
        |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::
Db      301 TTSVIOVGSFKLCRGCRIFRMAPIHHVELYHKGMPEPRVIVRFWIIISMLVLVLIGLATLKVR 360

```

```

RESULT 8
US-09-815-242-10028
; Sequence 10028, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011a
; CURRENT APPLICATION NUMBER: US/09/815.242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-03-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110

```

```
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10028
; LENGTH: 360
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-815-242-10028

Query Match          69.0%; Score 1268; DB 9; Length 360;
Best Local Similarity 66.4%; Pred. No. 1e-112;
Matches 239; Conservative 51; Mismatches 70; Indels 0; Gaps 0;

QY   1  ML LLLAEYLQQFVKFGVGYLTLRGILSVLTALSLSLWLGPMWMTLQIPQICQVRND 60
    :|::||::||::||::||::||::||::||::||::||::||::||::||::||::||
DB   1  MLVYLAHLVKYYSGFNVSFLTFRALIVSLTLFISLWMGPRWIAHLOKLSFGQVVVRND 60

QY   61 GPOSHLSKGTPTMGGALILTAITAISTILWADLSNRYVWVLVTLLFGAIGHVDDYRKV 120
    :|::||::||::||::||::||::||::||::||::||::||::||::||::||::||
DB   61 GPSHFHSKRGTPTMGGMIMLTALIVSILLWAYSPNYVMCVLVLYGVGYIGFVDYRKV 120

QY   121 IEKNSRGLPSRWKYFWQS VFGICAAVFLVWTATPIETTLIVPMLKSVEIQLGIFFFVLT 180
    :|::||::||::||::||::||::||::||::||::||::||::||::||::||::||
DB   121 VRDVTGKLARWKYFMSVIALGVAFALYAGDKTATQTQLVVPFFXDVMPQLGLFYILLA 180

QY   181 YFVIVGSSNAVNTDGLDGLAIIMPTVMVAGALGICYLSGNKVFKASYLIIPNVPGAGELI 240
    :|::||::||::||::||::||::||::||::||::||::||::||::||::||::||
DB   181 YFVIVGTGNAVNTDGLDGLAIIMPTVFVAGGFALVAWATCNMNFASYLHIPIYRHAGELY 240

QY   241 VFCAALVGAGLGFLENTYPAQVFMGDVGALALGAALTIAVIROEIVLFIMGGVFVME 300
    :|::||::||::||::||::||::||::||::||::||::||::||::||::||::||
DB   241 IVCTAIVAGLGFLENTYPAQVFMGDVGSLAGGALGIIVLRBGEFLVINGGVFVVE 300

QY   301 TLSVMIQVASFKLTGRRVFRMAPIHHEFLKGPDPDRVIVRFWIIITVLVILGLATLKLR 360
    :|::||::||::||::||::||::||::||::||::||::||::||::||::||::||
DB   301 TLSVIIQVSFKLRGQRIFRMAPIHHHYELKGWPPEPRVIVRFWIIISIMLVILGLATLKVR 360

RESULT 9
US-10-282-122A-56412
Sequence 56412, Application US/10282122A
Publication No. US20040029129A1
GENERAL INFORMATION:
APPLICANT: Wang, Liangsu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari
APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsythe, R.
APPLICANT: Xu, H.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: ELITEA 034A
CURRENT APPLICATION NUMBER: US/10/282,122A
CURRENT FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
```

```
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 56412
; LENGTH: 360
; TYPE: PR1
; ORGANISM: Escherichia coli
US-10-282-122A-56412

Query Match      59.0%; Score 1269; DB 12; Length 360;
Best Local Similarity 66.4%; Pred. No. 1e-112;
Matches 239; Conservative 51; Mismatches 70; Indels 0; Gaps 0;

QY 1 MLLLAELVLOQYKGFVGYOYLTLRGLSVLTALSLSLWLGPMWIRTLOIPQIQAVRND 60
DB 1 MLVWLAELVLYKYGFGNVFSLTFRAIVSLTLTALFISLWNGPRMIAHLQKLSFGQVRND 60
QY 61 GPQSHLSKGGPTMGGALILTAIATISLTLWADLSNRYVWVVLVTLFGLAGWDDYRKV 120
DB 61 GPESHFSKRGPTMGGMILTAVISVLLWAYPSNPYVWCVLVLYGVYGVFDDYRKV 120
QY 121 IEKNSRGLPSRWKYFQSGVFGIGAAVFLYMTAETPIETTLIVPMLKSVEIQIGFFVLT 180
DB 121 VRKOTGLIARKWYFWSVIALGVAFALYLAGKDTPTAQVLFVFFKDVMPQLGLFYLLA 180
QY 181 YFVIVGSSNAVNLTDGLDLAIMPTVMVAGALGIFCYLSGNVFAEYLIPNVFGAGELI 240
DB 181 YFVIVGTGNVNLTDGLDLAIMPTVFAAGFALVAVATGNMNFASFYHIFPYLRHAGELV 240
QY 241 VFCAALVAGLGFLEWNTYPAQVFMGVDGALAGALGTTAVIVROEIVLFGVGVFVME 300
DB 241 IVCTAIVGAGLGFLEWNTYPAQVFMGVDGSLAGLGGALGIIAVLLRQEFFLAVMGVGVFV 300
QY 301 TLSVMIQVAFPKLTGRVFRMAPHHHFELKGPDPDVIVRFRWIIITVILVIGLATLKL 360
DB 301 TLSVILQVGSFKLGRQIFRMAPHHHYELKGPDPDVIVRFRWIIISLMLVIGLATLKV 360

RESULT 10
US-10-282-122A-77920
; Sequence 77920, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: US/10/282,122A
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
```

```
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 77920
; LENGTH: 360
; TYPE: PR1
; ORGANISM: Yersinia pestis
US-10-282-122A-77920

Query Match      68.9%; Score 1265; DB 12; Length 360;
Best Local Similarity 65.6%; Pred. No. 2e-112;
Matches 236; Conservative 54; Mismatches 70; Indels 0; Gaps 0;

QY 1 MLLLAELVLOQYKGFVGYOYLTLRGLSVLTALSLSLWLGPMWIRTLOIPQIQAVRND 60
DB 1 MLVWLAELVLYKYGFGNVFSLTFRAIVSLTLTALFISLWNGPHLIAWLOKLQIQGVVRND 60
QY 61 GPQSHLSKGGPTMGGALILTAIATISLTLWADLSNRYVWVVLVTLFGLAGWDDYRKV 120
DB 61 GPESHFSKRGPTMGGMILFSTISVLMWAYPSNPYVWCVLVLYGVYGVFDDYRKV 120
QY 121 IEKNSRGLPSRWKYFQSGVFGIGAAVFLYMTAETPIETTLIVPMLKSVEIQIGFFVLT 180
DB 121 VRKNTKGLIARKWYFWSIIAALAAFTMYSIGKDTSALELVVFFKDIIMPQLGLLYLLA 180
QY 181 YFVIVGSSNAVNLTDGLDLAIMPTVMVAGALGIFCYLSGNVFAEYLIPNVFGAGELI 240
DB 181 YFVIVGTGNVNLTDGLDLAIMPTVFAAGFALVAVATGNMNFASFYHIFPYLRHAGELV 240
QY 241 VFCAALVAGLGFLEWNTYPAQVFMGVDGALAGALGTTAVIVROEIVLFGVGVFVME 300
DB 241 IVCTAIVGAGLGFLEWNTYPAQVFMGVDGSLAGLGGALGIIAVLLRQEFFLAVMGVGVFV 300
QY 301 TLSVMIQVAFPKLTGRVFRMAPHHHFELKGPDPDVIVRFRWIIITVILVIGLATLKL 360
DB 301 TLSVILQVGSFKLGRQIFRMAPHHHYELKGPDPDVIVRFRWIIISLMLVIGLATLKV 360

RESULT 11
US-10-282-122A-76018
; Sequence 76018, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: US/10/282,122A
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
```

PRIOR APPLICATION NUMBER: 60/207,727  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: 60/230,335  
PRIOR FILING DATE: 2000-09-06  
PRIOR APPLICATION NUMBER: 60/230,347  
PRIOR FILING DATE: 2000-09-09  
PRIOR APPLICATION NUMBER: 60/242,578  
PRIOR FILING DATE: 2000-10-23  
PRIOR APPLICATION NUMBER: 60/253,625  
PRIOR FILING DATE: 2000-11-27  
PRIOR APPLICATION NUMBER: 60/257,931  
PRIOR FILING DATE: 2000-12-22  
PRIOR APPLICATION NUMBER: 60/267,636  
PRIOR FILING DATE: 2001-02-09  
PRIOR APPLICATION NUMBER: 60/269,308  
PRIOR FILING DATE: 2001-02-16  
Remaining Prior Application data removed - See File Wrapper or PALM.  
NUMBER OF SEQ ID NOS: 78614  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 76018  
LENGTH: 360  
TYPE: PRT  
ORGANISM: Salmonella typhi  
US-10-282-122A-76018

Query Match 68.8%; Score 1264; DB 12; Length 360;  
Best Local Similarity 66.1%; Pred. No. 2.5e-112;  
Matches 238; Conservative 52; Mismatches 70; Indels 0; Gaps 0;  
Qy 1 MLLLAELVQFQYKGFVQYLTGLSILVLTALSLSLWLGPMWIRTLQIPQIGQVRND 60  
Db 1 MLVWLAHLVKYSGNFVSFLTFAIVSLTLTALFISLWMPRIARLQKLSFGQVRND 60  
Qy 61 GPQSHLSKKGTPMGAILTALTAISTLLWADLSNRYVWVLTLLFGAIGWDDYRKV 120  
Db 61 GPESHFSKRGTPMGIMLTALTAIVSVLLWAPSPNYVWCVLWVLLGYGIIGFVDDYRKV 120  
Qy 121 IEKNSRGLPSRWKYFWQSGVFGIGAAVFLYMTAETPIETTLIVPMLKSVIEIOLGIFVFLVLT 180  
Db 121 VRKDTKGLIARWKYFWMNVIALGVAFALYLVGKDTPTATQLVVFFKDVMPQLGLFYILLS 180  
Qy 181 YFVIVGSSNAVNLTDGLDLAIMPTVMVAGALGIFCYLSGNVKFAEYLLIPNVPAGELI 240  
Db 181 YFVIVGTGNVNLTDGLDLAIMPTVFAAGFALVAVATGNMNFANYLHPIYLRHAGELV 240  
Qy 241 VFCAALVAGLGLFLENTPYPAQVFMGDVGALGALGALGTIAIVROEIVLFIIMGVFWME 300  
Db 241 IVCTAIVGAGLGLFLENTPYPAQVFMGDVGSALGALGALGIIAIVLRQEBFLVIMGVFWVE 300  
Qy 301 TLSVMIQVASFKLTCRRVFRMAPIHHLHFKGWPDPRIIVRFWIIITVILVIGLATLKL 360  
Db 301 TLSVILQVGSFKLGRGIFRMAPIHHLHFKGWPDPRIIVRFWIIISLMLVIGLATLKL 360

RESULT 12  
US-10-282-122A-75186  
Sequence 75186, Application US/10282122A  
Publication No. US20040029129A1  
GENERAL INFORMATION:  
APPLICANT: Wang, Liangsu  
APPLICANT: Zamudio, Carlos  
APPLICANT: Malone, Cheryl  
APPLICANT: Haselbeck, Robert  
APPLICANT: Ohlsen, Kari  
APPLICANT: Zyskind, Judith  
APPLICANT: Wall, Daniel  
APPLICANT: Trawick, John  
APPLICANT: Carr, Grant  
APPLICANT: Yamamoto, Robert  
APPLICANT: Forsyth, R.  
APPLICANT: Xu, H.  
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
FILE REFERENCE: ELI/PA.034A

CURRENT APPLICATION NUMBER: US/10/282,122A  
CURRENT FILING DATE: 2003-02-20  
PRIOR APPLICATION NUMBER: 60/191,078  
PRIOR FILING DATE: 2000-03-21  
PRIOR APPLICATION NUMBER: 60/206,848  
PRIOR FILING DATE: 2000-05-23  
PRIOR APPLICATION NUMBER: 60/207,727  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: 60/230,335  
PRIOR FILING DATE: 2000-09-06  
PRIOR APPLICATION NUMBER: 60/230,347  
PRIOR FILING DATE: 2000-09-09  
PRIOR APPLICATION NUMBER: 60/242,578  
PRIOR FILING DATE: 2000-10-23  
PRIOR APPLICATION NUMBER: 60/253,625  
PRIOR FILING DATE: 2000-11-27  
PRIOR APPLICATION NUMBER: 60/257,931  
PRIOR FILING DATE: 2000-12-22  
PRIOR APPLICATION NUMBER: 60/267,636  
PRIOR FILING DATE: 2001-02-09  
PRIOR APPLICATION NUMBER: 60/269,308  
PRIOR FILING DATE: 2001-02-16  
Remaining Prior Application data removed - See File Wrapper or PALM.  
NUMBER OF SEQ ID NOS: 78614  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 75186  
LENGTH: 360  
TYPE: PRT  
ORGANISM: Salmonella typhimurium  
US-10-282-122A-75186

Query Match 68.6%; Score 1261; DB 12; Length 360;  
Best Local Similarity 65.8%; Pred. No. 4.7e-112;  
Matches 237; Conservative 53; Mismatches 70; Indels 0; Gaps 0;  
Qy 1 MLLLAELVQFQYKGFVQYLTGLSILVLTALSLSLWLGPMWIRTLQIPQIGQVRND 60  
Db 1 MLVWLAHLVKYSGNFVSFLTFAIVSLTLTALFISLWMPRIARLQKLSFGQVRND 60  
Qy 61 GPQSHLSKKGTPMGAILTALTAISTLLWADLSNRYVWVLTLLFGAIGWDDYRKV 120  
Db 61 GPESHFSKRGTPMGIMLTALTAIVSVLLWAPSPNYVWCVLWVLLGYGIIGFVDDYRKV 120  
Qy 121 IEKNSRGLPSRWKYFWQSGVFGIGAAVFLYMTAETPIETTLIVPMLKSVIEIOLGIFVFLVLT 180  
Db 121 VRKDTKGLIARWKYFWMNVIALGVAFALYLVGKDTPTATQLVVFFKDVMPQLGLFYILLS 180  
Qy 181 YFVIVGSSNAVNLTDGLDLAIMPTVMVAGALGIFCYLSGNVKFAEYLLIPNVPAGELI 240  
Db 181 YFVIVGTGNVNLTDGLDLAIMPTVFAAGFALVAVATGNMNFANYLHPIYLRHAGELV 240  
Qy 241 VFCAALVAGLGLFLENTPYPAQVFMGDVGALGALGALGTIAIVROEIVLFIIMGVFWME 300  
Db 241 IVCTAIVGAGLGLFLENTPYPAQVFMGDVGSALGALGALGIIAIVLRQEBFLVIMGVFWVE 300  
Qy 301 TLSVMIQVASFKLTCRRVFRMAPIHHLHFKGWPDPRIIVRFWIIITVILVIGLATLKL 360  
Db 301 TLSVILQVGSFKLGRGIFRMAPIHHLHFKGWPDPRIIVRFWIIISLMLVIGLATLKL 360

RESULT 13  
US-09-815-242-14075  
Sequence 14075, Application US/09815242  
Patent No. US20020061569A1  
GENERAL INFORMATION:  
APPLICANT: Haselbeck, Robert  
APPLICANT: Ohlsen, Kari L.  
APPLICANT: Zyskind, Judith W.  
APPLICANT: Wall, Daniel  
APPLICANT: Trawick, John D.  
APPLICANT: Carr, Grant J.  
APPLICANT: Yamamoto, Robert T.  
APPLICANT: Xu, H. Howard

```
/ TITLE OF INVENTION: Identification of Essential Genes in
; 
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11691
; LENGTH: 360
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(360)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-09-815-242-11691
```

Query Match      68.4%    Score 1257; DB 9; Length 360;  
Best Local Similarity    65.6%; Pred. No. 1.1e-111;  
Matches 236; Conservative 52; Mismatches 72; Indels 0; Gaps 0;

```
QY      1 MLLLAEYLQQFYKGFVQLTLRGILSVLTALSLSLWLGPMWRMTLPQTGAVRND 60  
DB      1 MLVLAHLVKYSYGFFVSFLTFRATISLVSLTALFISLMGPRIARLKLSFGQVRND 60  
  
QY     61 GPQHLSKKGTPTMGALILTATAISTILLWLADLSNRYVMVLVTLLFGAIGWDDYRKV 120  
DB     61 GPSHFSGKGTPTMGGIMILTAIVTSVLLWAYSPNYVCVLTVGLIGYGLIFGVDDIRKV 120  
  
QY    121 IEKNRGLPSRWKYFWQSDFVGIGAFAFLYMETAETPIETTLIVPMLKSVEIQGLIFFVLT 180  
DB    121 VRKDTKLARKWKYFMWSVIALGVAFALYLVGKDTPAQVPFPFKDWMPLGLFIYLLS 180  
  
QY    181 YFVIUGSSNAVNLTDGLDAIMPVVAGALGI FCYLSGNVKFAEYLIPNPVPGAGELI 240  
DB    181 YFVIUTGRANVLTDGLDLAIMPTVFVAAGPALVAMATNGNNFNANYLIHPYLRHAGELV 240  
  
QY    241 VFCAALVGAAGLFENWTYPQAQVFDGDVGCALAGAALGTIAVIRQEI VLFIMGVFEVME 300  
DB    241 IVCTAIVGAGLGFLENWTYPQAQVFDGDVGS LALGALGI IAVLRQEFLVINGGVFEV 300  
  
QY    301 TLSVMIQVASPKLTGRVRFRMAPIHHSFHFKGMPPDRPVIRVFWIITVLVLIGLATLKR 360  
DB    301 TLSVLQVGSFKLQGRIERMAPIIHHHYELKWPEPRVIRVFVIIISLMVLVIGLATLKR 360
```

RESULT 15  
US-10-282-122A-55841  
Sequence 55841, Application US/10282122A  
Publication No. US20040029129A1  
GENERAL INFORMATION:  
APPLICANT: Wang, Liangsu  
APPLICANT: Zamudio, Carlos  
APPLICANT: Malone, Cheryl  
APPLICANT: Haselbeck, Robert  
APPLICANT: Ohlsen, Kari  
APPLICANT: Zyskind, Judith  
APPLICANT: Wall, Daniel  
APPLICANT: Trawick, John  
APPLICANT: Carr, Grant  
APPLICANT: Yamamoto, Robert  
APPLICANT: Forsyth, R.  
APPLICANT: Xu, H.

```
/ TITLE OF INVENTION: Identification of Essential Genes in
; 
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14075
; LENGTH: 360
; TYPE: PRT
; ORGANISM: Salmonella typhi
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(360)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-09-815-242-14075
```

Query Match      68.5%    Score 1258; DB 9; Length 360;  
Best Local Similarity    65.8%; Pred. No. 9.2e-112;  
Matches 237; Conservative 52; Mismatches 71; Indels 0; Gaps 0;

```
QY      1 MLLLAEYLQQFYKGFVQLTLRGILSVLTALSLSLWLGPMWRMTLPQTGAVRND 60  
DB      1 MLVLAHLVKYSYGFFVSFLTFRATISLVSLTALFISLMGPRIARLKLSFGQVRND 60  
  
QY     61 GPQHLSKKGTPTMGALILTATAISTILLWLADLSNRYVMVLVTLLFGAIGWDDYRKV 120  
DB     61 GPSHFSGKGTPTMGGIMILTAIVTSVLLWAYSPNYVCVLTVGLIGYGLIFGVDDIRKV 120  
  
QY    121 IEKNRGLPSRWKYFWQSDFVGIGAFAFLYMETAETPIETTLIVPMLKSVEIQGLIFFVLT 180  
DB    121 VRKDTKLARKWKYFMWSVIALGVAFALYLVGKDTPAQVPFPFKDWMPLGLFIYLLS 180  
  
QY    181 YFVIUGSSNAVNLTDGLDAIMPVVAGALGI FCYLSGNVKFAEYLIPNPVPGAGELI 240  
DB    181 YFVIUTGRANVLTDGLDLAIMPTVFVAAGPALVAMATNGNNFNANYLIHPYLRHAGELV 240  
  
QY    241 VFCAALVGAAGLFENWTYPQAQVFDGDVGCALAGAALGTIAVIRQEI VLFIMGVFEVME 300  
DB    241 IVCTAIVGAGLGFLENWTYPQAQVFDGDVGS LALGALGI IAVLRQEFLVINGGVFEV 300  
  
QY    301 TLSVMIQVASPKLTGRVRFRMAPIHHSFHFKGMPPDRPVIRVFWIITVLVLIGLATLKR 360  
DB    301 TLSVLQVGSFKLQGRIERMAPIIHHHYELKWPEPRVIRVFVIIISLMVLVIGLATLXR 360
```

RESULT 14  
US-09-815-242-11691  
Sequence 11691, Application US/09815242  
Patent No. US20020061569A1  
GENERAL INFORMATION:  
APPLICANT: Haselbeck, Robert  
APPLICANT: Ohsen, Kari L.  
APPLICANT: Zyskind, Judith W.  
APPLICANT: Wall, Daniel  
APPLICANT: Trawick, John D.  
APPLICANT: Carr, Grant J.  
APPLICANT: Yamamoto, Robert T.  
APPLICANT: Yu, Howard

; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 55841
; LENGTH: 360
; TYPE: PRT
; ORGANISM: Enterobacter cloacae
US-10-282-122A-55841

Query Match	68.4%	Score 1257	DB 12	Length 360
Best Local Similarity	65.6%	Pred. No 1.1e-111		
Matches 236	Conservative 52	Mismatches 72	Indels 0	Gaps 0

  

QY	1	MLLLALYLOQFYKGFVQYILRLGILSVLTALSLMLGPMIIRTLQIPQIQAVRND	60
DB	1	MLVWLAHLVKYSGFNVSILTFRATVSLTALFISLWMPRIARLQKLSFGQWRND	60
QY	61	GPQSHLSKKGTPMGGALILTALTAISTLWADLSNRYVMVLAVTLLFGALGWDDYRKV	120
DB	61	GPESHFSKRGTPMGGIMILTATVSVLLWAFPSNPYVWCVLTVLVGYGIIGFVDDYRKV	120
QY	121	IEKNSRGLPSRWKIFYMQSVFGICAAVFLYMTAETPIETTLIVPMLKSVEIQIGIFFVVL	180
DB	121	VRKDTKGLIARWKYFWSVIALGVAFALYLAGKDTPATLVVPPFKDVMPLGLFYILLA	180
QY	181	YFVIVGSSNAVNLTDGLDGLAIMPVWAGALGIFCYLSGNVKFABYLLIPNVPAGELI	240
DB	181	YFVIVGTGNAVNLTDGLDGLAIMPVFAAGFALVAVATGNNMFANYLHPIYLRHAGELV	240
QY	241	VFCAALVGAAGLGFLENTPYPAQVFMGDVGALGALGALGTIAVIVROEIVLFIMGGVFYME	300
DB	241	IVCTAIVGAGLGFLENTPYPAQVFMGDVGSALGALGIIAVLRLQREFLLVIMGVSFVVE	300
QY	301	TLVSMITQVASFKLITGRVFRMAPIHHPHFKGWPPRIVRFWIITVILVIGLATLKL	360
DB	301	TLSVILQVGSFKLGRGRIFRMAPIHHPHFKGWPEPRIVRFWIISLMLVILGATLKV	360

Search completed: May 7, 2004, 08:46:23
Job time : 49 secs





Db 316 TLSVLQVGSFKLR-----RSAHLPH 336

RESULT 5

US-09-540-236-3138

; Sequence 3138, Application US/09540236

; Patent No. 6673910

; GENERAL INFORMATION:

; APPLICANT: Gary L. Breton et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATAR

; FILE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 2709.2005-001

; CURRENT APPLICATION NUMBER: US/09/540,236

; CURRENT FILING DATE: 2000-04-04

; NUMBER OF SEQ ID NOS: 3840

; SEQ ID NO 3138

; LENGTH: 368

; TYPE: PRT

; ORGANISM: M.catarrhalis

US-09-540-236-3138

Query Match 52.8%; Score 970.5; DB 4; Length 368;

Best Local Similarity 52.6%; Pred. No. 4.6e-101;

Matches 184; Conservative 69; Mismatches 86; Indels 11; Gaps 2;

QY 22 LTLRGLSVLTALSLSLWGPWMIRTLQIPQGOAVRNDGPQSHLSKKGTPMTMGALLIT 81

DB 19 LTLRALAVITSLIIIGKPVIRTLKYGQAVRNDGPKTLAKQGTPTMGGLVILV 78

QY 82 ATATSLWLADLSNRVYVWVLLFGAIGWVDDYRKVIEKNSRGLPSRWKYFWQSVFG 141

DB 79 ATGIATLWADLGNFVWILWVWIFGAVGADWDLKIKHKNPQGLARKKYFWLSVGS 138

QY 142 IGAALVYMTAETPIETTLI-----VPMKSV-----EIQGIPFVVLTVFVIGSSNA 190

DB 139 LFAGCSMYIAQTQDNTMIAMQDMLPLFKDIPFSAIPLGLGFIATYFVLGSSNA 198

QY 191 VNLTDGLDGLAMPVTVAGAGIFCYLSNVKAEYLLIPNVPAGELIVFCAALVAG 250

DB 199 VNLTDGLDGLVILFVVLVAGLGVFATISGPHFADYMHVPVIAVNAEVTIVCAAMIGAG 258

QY 251 LGFLWNTYPAQVFMGDVGALGALGTIAVIRQEIIVLFTMGVFMVMTLSVMIQVAS 310

DB 259 LGFLWNAAPADVFMGDVGALSJGGMLGTIAVMTQELAFAMGGLFVAEAVSVILQVGS 318

QY 311 FKLTGRVFRMAPIIHHFELKGPDPDRVIRVFWIITVLVILGLATLKL 360

DB 319 YRLKQRIIFMAPLHHHFEELGKETKVVARFYVICIILVILGLMTLKL 368

RESULT 6

US-09-134-001C-3851

; Sequence 3851, Application US/09134001C

; Patent No. 6380370

; GENERAL INFORMATION:

; APPLICANT: Lynn Doucette-Stamm et al

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS

; FILE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: GTC-007

; CURRENT APPLICATION NUMBER: US/09/134,001C

; CURRENT FILING DATE: 1998-08-13

; PRIOR APPLICATION NUMBER: US 60/064,964

; PRIOR FILING DATE: 1997-11-08

; PRIOR APPLICATION NUMBER: US 60/055,779

; PRIOR FILING DATE: 1997-08-14

; NUMBER OF SEQ ID NOS: 5674

; SEQ ID NO 3851

; LENGTH: 334

; TYPE: PRT

; ORGANISM: Staphylococcus epidermidis

US-09-134-001C-3851

Query Match 34.6%; Score 636; DB 4; Length 334;

Best Local Similarity 42.3%; Pred. No. 2.5e-63;

Matches 140; Conservative 70; Mismatches 93; Indels 28; Gaps 7;

QY 31 LTAISLSLWGPWMIRTLQIPQGOAVRNDGPQSHLSKKGTPMTMGALLITATAIST--- 87

DB 20 IIALITFILVPIILPTLRKMKFGQIRBEGPOSHMKKTGTPMGGTTFIISIIIIIIIA 79

QY 88 LNWADLSNRVYVWVLLFGAIGWVDDYRKVIEKNSRGLPSRWKYFWQSVFGICAAVF 147

DB 80 IIFVDHSPNII-LLLFTVIGFGLIGFDDYIIIVKXNQLTSKQKFLAQIIII---AVI 134

QY 148 LYMTAE-----TPIETTLIVPMKLSVBIQLGIFPVVTVFYVIGSSNAVNLTDGLDGLAIM 203

DB 135 FEVLSDVFEHLVHFTDHLHPFV-NFDIPLSFAYVIFVFMQVGFSAVNLTDGLDGLAT- 192

QY 204 PTVMVAGALGIFCYLSNVKAEYLLIPNVPAGELIVFCAALVAGAGLGFELWNTYPAQV 263

DB 193 -----GLSI-----IGFAMVMSYMLDSPAIGFICIIINFALLGFLPYNLNPAKV 238

QY 264 FMGDVGALGALGALGTIAVIRQEIIVLFTMGVFMVMTLSVMIQVASFKLTGRVFRMAP 323

DB 239 FMGDTGSLAGGIFATISIMLNQELSILIGFVVFVETLSVMLQVASYKLTKKRIFKMS 298

QY 324 IHHFELKGPDPDRVIRVFWIITVLVILGL 354

DB 299 IHHFELSNGWGWKVVTVFWVGLITGLIGL 329

RESULT 7

US-09-107-532A-5276

; Sequence 5276, Application US/09107532A

; Patent No. 6583275

; GENERAL INFORMATION:

; APPLICANT: Lynn A Doucette-Stamm and David Bush

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO

; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS

; NUMBER OF SEQUENCES: 7310

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: GENOME THERAPEUTICS CORPORATION

; STREET: 100 Beaver Street

; CITY: Waltham

; STATE: Massachusetts

; COUNTRY: USA

; ZIP: 02354

; COMPUTER READABLE FORM:

; MEDIUM TYPE: CD-ROM ISO9660

; COMPUTER: PC

; OPERATING SYSTEM: <Unknown>

; SOFTWARE: ASCII

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/107,532A

; FILING DATE: 30-Jun-1998

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 60/085,598

; FILING DATE: 14 May 1998

; APPLICATION NUMBER: 60/051571

; FILING DATE: July 2, 1997

; ATTORNEY/AGENT INFORMATION:

; NAME: Ariniello, Pamela Deneke

; REGISTRATION NUMBER: 40,489

; REFERENCE/DOCKET NUMBER: GTC-012

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (781)893-5007

; TELEFAX: (781)893-8277

; INFORMATION FOR SEQ ID NO: 5276:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 331 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; HYPOTHETICAL: YES

; ORIGINAL SOURCE:



; NUMBER OF SEQUENCES: 4  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Eli Lilly and Company  
 ; STREET: Lilly Corporate Center  
 ; CITY: Indianapolis  
 ; STATE: Indiana  
 ; COUNTRY: US  
 ; ZIP: 46285  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patentin Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/986,768  
 ; FILING DATE: December 8, 1997  
 ; CLASSIFICATION: 424  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Webster, Thomas D  
 ; REGISTRATION NUMBER: 39,872  
 ; REFERENCE/DOCKET NUMBER: X-11753  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 317/276-3334  
 ; INFORMATION FOR SEQ ID NO: 2:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 326 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 ; US-08-986-768-2

Query Match 28.3%; Score 520; DB 3; Length 326;  
 Best Local Similarity 36.9%; Pred. No. 3.1e-50;  
 Matches 129; Conservative 60; Mismatches 111; Indels 50; Gaps 9;  
 QY 22 LTRGLSVLTALSLSLWGPWIRTLQIPQIQAVERNNDGP-QSHLSKKGTPMTGGALIL 80  
 DB 14 LTLVGI-----PAFIQYRKAQITGOQHEDVKQ-HQAKAGTPTMGSLVFL 58  
 QY 81 TATAISTLLWADLSNEY---VMVLVVTLLFGAIGWDDYKRVIEKNSRGLPSRWKYFWQ 137  
 DB 59 ITSVLVAFPAFFLSQFSNNVGMILFILVYGLVFLDQFLVKFKINEGLNPKQKLAQ 118  
 QY 138 SVFGIGAAVFLYMTAETPIETTLIVPMLKSVBIQLGIFVFWLTYRVIVGSSNAVNLTDGL 197  
 DB 119 LLGGV--IFVLVFERGGDILSVFGYP-----VHLGFFYIPFALFWLVGFSNAVNLTDGV 170  
 QY 198 DGLAIMPVTVAGALGIFCYLSGNVKAEBYLLIPNVPGAGELIVFCAALVGAGLGLFWFN 257  
 DB 171 DGLASISVVISLFAYGVIAVQGM-----DILLVILAMIGLLGFFIFN 215  
 QY 258 TYPAQVFMGDVGALALGAALGTIAVTRQEIIVLFIMGGVFMVMTLSVMIQVASFKLT-GR 316  
 DB 216 HKPAKVFMGDVGSALGGMALASMALHQAETLLIIGIVYVFTTSVMVQVSFKLTGK 275  
 QY 317 RVFRMAPIHHPHKLK-----WDPDRVIVFWITVILVILGLATLKL 359  
 DB 276 RIFRMTFVHHFELGLSLGKGNFWSKVDFFFWGVGLLASLTLLAILYL 325

RESULT 11  
 US-09-107-532A-4764  
 ; Sequence 4764, Application US/09107532A  
 ; Patent No. 6583275  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Lynn A Doucette-Stamm and David Bush  
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
 ; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS  
 ; NUMBER OF SEQUENCES: 7310  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: GENOME THERAPEUTICS CORPORATION  
 ; STREET: 100 Beaver Street  
 ; CITY: Walham

; STATE: Massachusetts  
 ; COUNTRY: USA  
 ; ZIP: 02354  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: CD/ROM ISO9660  
 ; COMPUTER: PC  
 ; OPERATING SYSTEM: <Unknown>  
 ; SOFTWARE: ASCII  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/107,532A  
 ; FILING DATE: 30-Jun-1998  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 60/085,598  
 ; FILING DATE: 14 May 1998  
 ; APPLICATION NUMBER: 60/051571  
 ; FILING DATE: July 2, 1997  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Ariniello, Pamela Deneke  
 ; REGISTRATION NUMBER: 40,489  
 ; REFERENCE/DOCKET NUMBER: GTC-012  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (781)893-5007  
 ; TELEFAX: (781)893-8277  
 ; INFORMATION FOR SEQ ID NO: 4764:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 386 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: Protein  
 ; HYPOTHETICAL: YES  
 ; ORIGINAL SOURCE:  
 ; ORGANISM: Enterococcus faecium  
 ; FEATURE:  
 ; NAME/KEY: misc feature  
 ; LOCATION: (B) LOCATION 1...386  
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 4764:  
 US-09-107-532A-4764

Query Match 13.3%; Score 245; DB 4; Length 386;  
 Best Local Similarity 26.6%; Pred. No. 4.9e-19;  
 Matches 101; Conservative 65; Mismatches 131; Indels 82; Gaps 17;  
 QY 22 LTRGLSVLTALSLSLWGPWIRTLQIPQIQAVERNNDGP-QSHLSKKGTPMTGGALIL 80  
 DB 15 MVIRFFETIL---LSLITLP-VFKIISV-QTGMV---DKPNERRINKVMPSPAGGLPIF 65  
 QY 81 TATAISTLLW--ADLSNRYVMVLVVTLLFGAIGWDDYKRVIEKNSRGLPSRWKYFWQS 138  
 DB 66 ISFVISTLFLFKLIIPQSYIVPILIASLVIITVGLDDKYELTPKQ-----XS 113  
 QY 139 VFGIGAAVFLYMTAETPIETTLIVPMLKSVBIQLGIFVFWLTYRVIVGSSNAVNLTDGLD 198  
 DB 114 IGLIASLIIFYVADIRID-SFTLPFIG--YIQLGWLSPFVTFWIFGITNAVNLIDGLD 170  
 QY 199 GLAIMPVTVAGALGIFCYLSGNVKAEBYLLIPNVPGAGELIVFCAALVGAGLGLFWFN 258  
 DB 171 GLAIGISLIGLITIGIYF---FLHASTYIP-----VWIFC--LVASIIIGFFPNF 218  
 QY 259 YPAQVFMGDVGALALGAALGTIA-----VIVRQEIIVLFIMGGVFMVMTLSVMIQVASFK 312  
 DB 219 YPAKIVLGDGTGALFGPMMAVLSLQGLKNVTFVSSISLLIIVMGVPVTDTPFAIR-----R 274  
 QY 313 LTGRVFRMAP---IHHPFLKGMPPDRVIVRW-----IITVILV- 350  
 DB 275 KANRVSFSTADKKHLHRLHLSLGTHTKGAVLTIYAMALMFSFTAMVWNYTGRIGTILII 334  
 QY 351 -----LIGLATLK 358  
 DB 335 AMLFAAILLPELIGLINEK 353

RESULT 12  
 US-09-634-238-311

```

; Sequence 311, Application US/09634238
; Patent No. 6544772
; GENERAL INFORMATION:
; APPLICANT: Glenn, Matthew
; APPLICANT: Havukkala, Ilkka J.
; APPLICANT: Bicksberg, Leonard, N.
; APPLICANT: Lubbers, Mark W.
; APPLICANT: Dekker, James
; APPLICANT: Christenson, Anna C.
; APPLICANT: Holland, Rose
; APPLICANT: O'Toole, Paul W.
; APPLICANT: Reid, Julian R.
; APPLICANT: Coolbear, Timothy
; TITLE OF INVENTION: Polynucleotides
; TITLE OF INVENTION: them and method
; FILE REFERENCE: 11000.1043U1
; CURRENT APPLICATION NUMBER: US/09/63
; CURRENT FILING DATE: 2000-06-08
; NUMBER OF SEQ ID NOS: 422
; SOFTWARE: FastSeq for Windows Versio
; SEQ ID NO 311
; LENGTH: 106
; TYPE: PRT
; ORGANISM: Lactobacillus rhamnosus
US-09-634-238-311

```

Query Match	12.6%;	Score 231.5;	DB 4;	Length 106;
Best Local Similarity	45.5%;	Pred. No. 2.3e-18;		
Matches	56;	Conservative 14;	Mismatches 36;	Indels 17; Gaps 3;
Qy	159	TLIVPMLKSYEIQIGIPVVLTYFVIVGSSNAVNLTDGLDGLAIPPTVMVAGALGIFCYL	218	
Db	1	TLWVPVLGNVATW--FVIVFVFWLVGFSNAVNLTDGLDGL-----VAG-----	43	
Qy	219	SGNVKFAEYLLIPNVPGAGELIVFCALVGAGLGFLEMTYPAQVFMGDVGALGAALG	278	
Db	44	QTTISFATYAVIAREGEVDVLLICLVTVGMLGFLFLENHKPAQIFMGDLSLGGMLA	103	
Qy	279	TIA	281	
Db	104	VAA	106	

```

RESULT 13
US-09-134-001C-5489
; Sequence 5489, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: LYNN Douchette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND NUCLEIC ACID SEQUENCES
; TITLE OF INVENTION: EPIDERMIDIS FOR D
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,96
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,77
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 5489
; LENGTH: 375
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-5489

```

```

Query Match      12.1%; Score 222.5; DB 4; Length 375;
Best Local Similarity 25.0%; Pred. No. 1.6e-16;
Matches 77; Conservative 51; Mismatches 103; Indels 77; Gaps 10;

QY 69 KGTPTMGAGALILTAISTALWADLSNRYVMVLVTLLFGAI-----GWYDD-----Y 117
db 58 KPISVMGVTIIFSLIGTWGHPTE-----VKPLILGAIITMVMVGLDIDYDLRPY 111

```

Qy	118	RVIEKNSRGLPSRWKYFWQSVFGIGAAVFLYMTAET-----PIETTLIVPMLKSVIEIQ	171
Db	112	LKLAGQ-----IVAALIVTFYGITIDFISLPIGPT-----IH	143
Qy	172	LGTFPVVLTFVIVGSSNAVNLTDGLDGLAIMPTVMVAGALGFCYLSGNVKFAEYLLIP	231
Db	144	PGFSFITWIVWIVAININLIDGLDGLASGVSAALMTIGFTAILQANI-----	194
Qy	232	NVPGAGELIVFCALVGAGIGFLFWNTYPAQVFMGVDYGMALGAALGTIAVIVRQEIVL-	290
Db	195	-----FIIMCCVLLGSLGLGFYFNHPKAFIFGLDGLMIGFIIGFLSULGFKNITFI	248
Qy	291	-----FTMGGVFMVMTLSVMIQVASPFLKTGRVRFR--MAPIHHHPELKGWPPDPVIVRFPW	343
Db	249	ALFPPIVILAVPFDITLFIATIRMKM---KGOHIMQADKSHLHKILALGYTHRQIVLLIY	305
Qy	344	IITVILVL	351
Db	306	SIAMFSL	313

RESULT 14  
US-09-489-039A-12452  
; Sequence 12452, Application US/09489039A  
; Patent NO. 5610836  
; GENERAL INFORMATION:  
; APPLICANT: Gary Breton et. al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA  
; TITLE OF INVENTION: PNEUMONIA FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 2709.2004001  
; CURRENT APPLICATION NUMBER: US/09/489,039A  
; CURRENT FILING DATE: 2000-01-27  
; PRIOR APPLICATION NUMBER: US 60/117,747  
; PRIOR FILING DATE: 1999-01-29  
; NUMBER OF SEQ ID NOS: 14342  
; SEQ ID NO 12452  
; LENGTH: 69  
; TYPE: PRT  
; ORGANISM: Klebsiella pneumoniae  
US-09-489-039A-12452

Query Match	12.1%	Score 222	DB 4	Length 69
Best Local Similarity	75.0%	Pred. No. 1.5e-17		
Matches 39	Conservative 8	Mismatches 5	Indels 0	Gaps 0

  

QY	309	ASFKLTSRVERFMAPIHHHPKLGWPPRVIVRFWIIIVILGLATLKL	360
	1		
	10		

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS

. TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS

RESULT 15  
 US-08-846-762-92  
 ; Sequence 92, Application US/08846762A  
 ; Patent No. 5994072  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Lam, Joseph S.  
 ; APPLICANT: Burrows, Lori  
 ; APPLICANT: Charter, Deborah  
 ; APPLICANT: de Kievit, Teresa  
 ; TITLE OF INVENTION: No. 5994072a1 Proteins Involved in the Synthesis and Assembly  
 ; TITLE OF INVENTION: of C-Antigen in Pseudomonas Aeruginosa  
 ; FILE REFERENCE: 6580-089  
 ; CURRENT APPLICATION NUMBER: US/08/846,762A  
 ; CURRENT FILING DATE: 1997-04-30  
 ; NUMBER OF SEQ ID NOS: 100  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 92  
 ; LENGTH: 341  
 ; TYPE: prt  
 ; ORGANISM: Pseudomonas aeruginosa  
 ; US-08-846-762-92

Query Match  
10.5%; Score 192; DB 2; Length 341;

Best Local Similarity 24.7%; Pred. No. 4e-13;  
 Matches 83; Conservative 57; Mismatches 116; Indels 80; Gaps 16;

```

Qy 19 FOYLTGILSVLTALSLSLWLGPMWIRTLOIPQIGQAVRN---DGQSHLSK-KGTPTM 74
Db 5 FELLTFELLSVGLTYLRLY-----ALKNNIIDTNSRSSHVTPTPRG 48

Qy 75 GGALILTAIAISTLMADLSNRYVWVLVVTLL-----FGAIGWDDYRKVIEKNSRGLP 129
Db 49 GGVAIVISFLIGIILFYFLG--YLPILSVGLIVSGGVIALVGFWDHGHIAAR----- 100

Qy 130 SRWKYFQSVFVGIGAAVF-LYMTAETPIETTLIVPMLKSVEIQIGIPVVLTYFVIYVSS 188
Db 101 -----WLLAHFSAAAFLLFCFGFPVLN-----VSGFIIEIGFGLFLVWML 148

Qy 189 NAVNLTDGLDGLAIMPTVMVA-GALGIFCYLSG-NVKEAYLLIPNVFEGAGELIVFCAAL 246
Db 149 NLYNFMDCIGDGLASAEAVTACIGMTAIY-YISGDHIELNSFLV-----WLLACTV- 198

Qy 247 VGAGLGLFWNTYPAQVFMGDVGALALGALGTIAVI-----VRQEIYLFIMGVFWMET 301
Db 199 ----LGFLWNFPFPAKIFMGDAGSGFLGLMIGSLAISAGWIDTRFFCWLILLGLFIVDA 254

Qy 302 LSVMIQ--VASFK-----LTGRRVFRMAPI 324
Db 255 TWLVRRVVGFKVYEHRSHGYQIASRRFKHLPV 290

```

Search completed: May 7, 2004, 08:41:31  
 Job time : 24 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2004, 08:37:37 ; Search time 59 Seconds  
(without alignments)  
1724.019 Million cell updates/sec

Title: US-10-089-787-2  
Perfect score: 1837  
Sequence: 1 MLLLLAALYQFYKPGVQF.....RFTITVLVLIQGLATKLKLR 360

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A Geneseq\_29Jan04:\*

- 1: geneseqp1980s:\*
- 2: geneseqp1990s:\*
- 3: geneseqp2000s:\*
- 4: geneseqp2001s:\*
- 5: geneseqp2002s:\*
- 6: geneseqp2003as:\*
- 7: geneseqp2003bs:\*
- 8: geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1837	100.0	360	4 AAB73487	Aab73487 Pseudomon
2	1824	99.3	360	4 AAU36418	Aau36418 Pseudomon
3	1824	99.3	360	6 ABU38729	Abu38729 Protein e
4	1689	91.9	360	6 ABU39577	Abu39577 Protein e
5	1658	90.3	360	6 ABU41789	Abu41789 Protein e
6	1822	70.3	360	6 ABU49535	Abu49535 Protein e
7	1280	69.7	361	6 ABM68871	Abm68871 Photorhab
8	1274	69.4	360	6 ABU38972	Abu38972 Protein e
9	1268	69.0	360	4 AAU34435	Aau34435 E. coli c
10	1268	69.0	360	4 AAG98403	Aag98403 Escherich
11	1268	69.0	360	6 ABU28488	Abu28488 Protein e
12	1265	68.9	360	6 ABU49896	Abu49896 Protein e
13	1264	68.8	360	6 ABU48094	Abu48094 Protein e
14	1261	68.6	360	6 ABU47262	Abu47262 Protein e
15	1258	68.5	360	4 AAU38482	Aau38482 Salmonell
16	1257	68.4	360	4 AAU36098	Aau36098 Klebsiell
17	1257	68.4	360	6 ABU27917	Abu27917 Protein e
18	1257	68.4	360	6 ABU31662	Abu31662 Protein e
19	1250	68.0	360	6 ABU40890	Abu40890 Protein e
20	1228	66.8	352	6 ABU45309	Abu45309 Protein e
21	1216	66.2	360	4 AAU35593	Aau35593 Haemophil
22	1216	66.2	360	6 ABU30495	Abu30495 Protein e
23	1205.5	65.6	389	6 ABU22089	Abu22089 Protein e
24	1202.5	65.5	389	6 ABU23453	Abu23453 Protein e
25	1185.5	64.5	389	6 ABU21181	Abu21181 Protein e

ALIGNMENTS

RESULT 1  
AAB73487  
ID AAB73487 standard; protein; 360 AA.  
XX  
AC AAB73487;  
XX  
17-JUL-2001 (first entry)  
XX  
DE Pseudomonas aeruginosa Mray protein.  
XX  
KW Mray; phospho-N-acetylmuramoyl-pentapeptide translocase;  
KW peptidoglycan biosynthesis; bacterial cell wall; infection;  
KW drug screening; antibacterial; Pseudomonad; G+C rich bacterium.  
XX  
OS Pseudomonas aeruginosa.  
XX  
WO200125251-A1.  
XX  
12-APR-2001.  
XX  
29-SEP-2000; 2000WO-US027056.  
XX  
04-OCT-1999; 99US-0157580P.  
XX  
(MERI ) MERCK & CO INC.  
XX  
El-Sherbeini M, Azzolina B;  
XX  
WPI; 2001-308221/32.  
XX  
N-PSDB; AAP86639.  
XX  
New Mray gene and enzyme of Pseudomonas aeruginosa, useful in vitro  
assays for screening antibacterial compounds that target cell wall  
biosynthesis, particularly for screening antibiotics against  
Pseudomonads.  
XX  
Claim 15; Fig 1; 22pp; English.  
XX  
This sequence represents the Pseudomonas aeruginosa phospho-N-  
acetylmuramoyl-pentapeptide translocase, designated Mray. The Mray  
protein is involved in bacterial cell wall biosynthesis, it catalyses the  
first step of the membrane cycle of peptidoglycan biosynthesis, i.e., the  
transfer of an N-acetylmuramic acid peptide to a bactoprenol phosphate  
carrier molecule. The Mray protein is useful in in vitro assays to screen  
for antibacterial compounds that target cell wall biosynthesis.  
XX  
Inhibitors of the Mray protein are useful in preventing the growth of  
Pseudomonads and other G+C rich bacteria. Pseudomonas aeruginosa is an  
opportunistic pathogen which causes infections in patients with burns,

26	1118	60.9	372	6	ABU16792	Abu16792 Protein e
27	1118	60.9	372	6	ADA34882	Ada34882 Acinetoba
28	1099	59.8	360	6	ABU38140	Abu38140 Protein e
29	1099	59.8	374	3	AAU74406	Aay74406 Neisseria
30	1094	59.6	360	6	ABP80568	Abp80568 N. gonorr
31	1094	59.6	360	6	ABU37206	Abu37206 Protein e
32	1069	58.2	376	3	AAU74404	Aay74404 Neisseria
33	972.5	52.9	365	6	ABU35134	Abu35134 Protein e
34	901.5	49.1	239	3	ABU33543	Abu33543 Protein e
35	893.5	48.6	1286	4	ABG25629	Abg25629 Novel hum
36	893.5	48.6	1286	4	ABG25994	Abg25994 Novel hum
37	893.5	48.6	1286	4	ABG25086	Abg25086 Novel hum
38	812.5	44.2	351	6	ABU19161	Abu19161 Protein e
39	782.5	42.6	353	6	ABU26395	Abu26395 Protein e
40	755.5	41.1	215	6	AAU20164	Aau20164 Protein e
41	709.5	38.6	376	2	AAW20932	Aaw20932 H. pylori
42	708.5	38.6	353	4	AAU35735	Aau35735 Helicobac
43	708.5	38.6	353	6	ABU30794	Abu30794 Protein e
44	682	37.1	376	3	AAU74405	Aay74405 Neisseria
45	671	36.5	324	6	ABU18859	Abu18859 Protein e

CC neutropenia, or cystic fibrosis. Primers or probes derived from the may  
CC gene are useful in nucleic acid amplification-based assays for detecting  
CC the presence of a polynucleotide encoding *Pseudomonas aeruginosa* May  
CC protein  
XX  
SQ Sequence 360 AA;  
  
Query Match 100.0%; Score 1837; DB 4; Length 360;  
Best Local Similarity 100.0%; Pred. No. 2.3e-193;  
Matches 360; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 MLLLLAEYLQOYKGFYQVLTGILSVLTALSLSLWLGPMWIRTLOIPQIGQVRND 60  
DB 1 MLLLLAEYLQOYKGFYQVLTGILSVLTALSLSLWLGPMWIRTLOIPQIGQVRND 60  
  
QY 61 GPQSHLSKGGTPTMGAGLILTAISTLLWADLSNRYVWVLTLLFGAIGWDDYRKV 120  
DB 61 GPQSHLSKGGTPTMGAGLILTAISTLLWADLSNRYVWVLTLLFGAIGWDDYRKV 120  
  
QY 121 IEKNSRGLPSRWKYPWQSVFGIGAAVFLYMTAETPIETTLIVPMLKSVRIQLGIFVVL 180  
DB 121 IEKNSRGLPSRWKYPWQSVFGIGAAVFLYMTAETPIETTLIVPMLKSVRIQLGIFVVL 180  
  
QY 181 YFVIVGSSNAVNLTDGLDLAIMPTVMVAGALGIFCYLSGNVKFAEYLLIPNVPAGAGELI 240  
DB 181 YFVIVGSSNAVNLTDGLDLAIMPTVMVAGALGIFCYLSGNVKFAEYLLIPNVPAGAGELI 240  
  
QY 241 VFCAALVGAGLGLFWNTYPAQVFMGVDVGALGALGTTTAVIVROEIVLFGGVFVME 300  
DB 241 VFCAALVGAGLGLFWNTYPAQVFMGVDVGALGALGTTTAVIVROEIVLFGGVFVME 300  
  
QY 301 TLSVMIQVASFKLGTGRVFRFMAPIHFFHFKLWGPDPFVIVRFWIIITVILVILGLATLKL 360  
DB 301 TLSVMIQVASFKLGTGRVFRFMAPIHFFHFKLWGPDPFVIVRFWIIITVILVILGLATLKL 360  
  
RESULT 2  
AAU36418  
ID AAU36418 standard; protein; 360 AA.  
XX  
AC AAU36418;  
XX  
DT 14-FEB-2002 (first entry)  
DE  
XX  
KW *Pseudomonas aeruginosa* cellular proliferation protein #408.  
XX  
KW Antisense; prokaryotic cellular proliferation protein; antibiotic;  
KW antibacterial; drug design.  
OS *Pseudomonas aeruginosa*.  
XX  
FN WO200170955-A2.  
XX  
PD 27-SEP-2001.  
XX  
PF 21-MAR-2001; 2001WO-US009180.  
XX  
PR 21-MAR-2000; 2000US-0191078P.  
PR 23-MAY-2000; 2000US-0206848P.  
PR 26-MAY-2000; 2000US-0207727P.  
PR 23-OCT-2000; 2000US-0242578P.  
PR 27-NOV-2000; 2000US-0253625P.  
PR 22-DEC-2000; 2000US-0257931P.  
PR 16-FEB-2001; 2001US-0269308P.  
XX  
FA (ELIT-) ELITRA PHARM INC.  
XX  
PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;  
PI Yamamoto RT, Xu HH;  
XX  
XX WPI; 2001-611495/70.  
DR N-PSDB; AAS54277.  
XX

PT New polynucleotides for the identification and development of  
PT antibiotics, comprise sequences of antisense nucleic acids.  
XX  
PS Example 3; SEQ ID NO 12011; 511pp; English.  
XX  
CC The invention relates to antisense inhibitors of genes essential to  
CC prokaryotic cellular proliferation, their use in identifying the genes,  
CC their use in the discovery of novel antibiotics, the essential genes  
CC themselves and the encoded proteins. The prokaryotes used are *Escherichia*  
CC coli, *Staphylococcus aureus*, *Salmonella typhi*, *Klebsiella pneumoniae*,  
CC *Pseudomonas aeruginosa* and *Enterococcus faecalis*. The invention is also  
CC useful for the identification of potential new targets for antibiotic  
CC development. The antisense nucleic acids can also be used to identify  
CC proteins used in proliferation, to express these proteins, and to obtain  
CC antibodies capable of binding to the expressed proteins. The proteins can  
CC be used to screen compounds in rational drug discovery programmes. The  
CC antisense nucleic acid sequence is also useful to screen for homologous  
CC nucleic acids which are required for cell proliferation in a wide variety  
CC of organisms. The present sequence represents an essential prokaryotic  
CC cellular proliferation protein. Note: The sequence data for this patent  
CC did not form part of the printed specification, but was obtained in  
CC electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 360 AA;  
  
Query Match 99.3%; Score 1824; DB 4; Length 360;  
Best Local Similarity 99.4%; Pred. No. 6.1e-192;  
Matches 358; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
  
QY 1 MLLLLAEYLQOYKGFYQVLTGILSVLTALSLSLWLGPMWIRTLOIPQIGQVRND 60  
DB 1 MLLLLAEYLQOYKGFYQVLTGILSVLTALSLSLWLGPMWIRTLOIPQIGQVRND 60  
  
QY 61 GPQSHLSKGGTPTMGAGLILTAISTLLWADLSNRYVWVLTLLFGAIGWDDYRKV 120  
DB 61 GPQSHLSKGGTPTMGAGLILTAISTLLWADLSNRYVWVLTLLFGAIGWDDYRKV 120  
  
QY 121 IEKNSRGLPSRWKYPWQSVFGIGAAVFLYMTAETPIETTLIVPMLKSVRIQLGIFVVL 180  
DB 121 IEKNSRGLPSRWKYPWQSVFGIGAAVFLYMTAETPIETTLIVPMLKSVRIQLGIFVVL 180  
  
QY 181 YFVIVGSSNAVNLTDGLDLAIMPTVMVAGALGIFCYLSGNVKFAEYLLIPNVPAGAGELI 240  
DB 181 YFVIVGSSNAVNLTDGLDLAIMPTVMVAGALGIFCYLSGNVKFAEYLLIPNVPAGAGELI 240  
  
QY 241 VFCAALVGAGLGLFWNTYPAQVFMGVDVGALGALGTTTAVIVROEIVLFGGVFVME 300  
DB 241 VFCAALVGAGLGLFWNTYPAQVFMGVDVGALGALGTTTAVIVROEIVLFGGVFVME 300  
  
QY 301 TLSVMIQVASFKLGTGRVFRFMAPIHFFHFKLWGPDPFVIVRFWIIITVILVILGLATLKL 360  
DB 301 TLSVMIQVASFKLGTGRVFRFMAPIHFFHFKLWGPDPFVIVRFWIIITVILVILGLATLKL 360  
  
RESULT 3  
ABU38729  
ID ABU38729 standard; protein; 360 AA.  
XX  
AC ABU38729;  
XX  
DT 19-JUN-2003 (first entry)  
DE  
XX  
DE Protein encoded by Prokaryotic essential gene #24256.  
XX  
KW Antisense; prokaryotic essential gene; cell proliferation; drug design.  
XX  
OS *Pseudomonas aeruginosa*.  
XX  
PN WO200277183-A2.  
XX  
XX 03-OCT-2002.  
XX

PF 21-MAR-2002; 2002WO-US009107.  
 PR 21-MAR-2001; 2001US-00815242.  
 PR 06-SEP-2001; 2001US-00948993.  
 PR 25-OCT-2001; 2001US-0342923P.  
 PR 08-FEB-2002; 2002US-00072851.  
 PR 06-MAR-2002; 2002US-0362699P.  
 XX (ELIT-) ELITRA PHARM INC.  
 XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;  
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;  
 XX WPI; 2003-029926/02.  
 DR N-PSDB; ACA42599.  
 XX New antisense nucleic acids, useful for identifying proteins or screening  
 PT for homologous nucleic acids required for cellular proliferation to  
 PT isolate candidate molecules for rational drug discovery programs.  
 XX Claim 25; SEQ ID NO 66653; 1766pp; English.  
 XX The invention relates to an isolated nucleic acid comprising any one of  
 CC the 6213 antisense sequences given in the specification where expression  
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:  
 CC (1) a vector comprising a promoter operably linked to the nucleic acid  
 CC encoding a polypeptide whose expression is inhibited by the antisense  
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated  
 CC polypeptide or its fragment whose expression is inhibited by the  
 CC antisense nucleic acid; (4) an antibody capable of specifically binding  
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular  
 CC proliferation or the activity of a gene in an operon required for  
 CC proliferation; (7) identifying a compound that influences the activity of  
 CC the gene product or that has an activity against a biological pathway  
 CC required for proliferation, or that inhibits cellular proliferation; (8)  
 CC identifying a gene required for cellular proliferation or the biological  
 CC pathway in which a proliferation-required gene or its gene product lies  
 CC or a gene on which the test compound that inhibits proliferation of an  
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a  
 CC compound's activity; (11) a culture comprising strains in which the gene  
 CC product is overexpressed or underexpressed; (12) determining the extent  
 CC to which each of the strains is present in a culture or collection of  
 CC strains; or (13) identifying the target of a compound that inhibits the  
 CC proliferation of an organism. The antisense nucleic acids are useful for  
 CC identifying proteins or screening for homologous nucleic acids required  
 CC for cellular proliferation to isolate candidate molecules for rational  
 CC drug discovery programs, or for screening homologous nucleic acids  
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,  
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of  
 CC the target prokaryotic essential genes. Note: The sequence data for this  
 CC patent did not form part of the printed specification, but was obtained  
 CC in electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 XX SQ Sequence 360 AA;  
 Query Match 99.3%; Score 1824; DB 6; Length 360;  
 Best Local Similarity 99.4%; Pred. No. 6.1e-192;  
 Matches 358; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 MLLLLAEYLQYKFGVGYLTLRGLSVLTSLSLWLGPMWIRTLQIQIGQAVRND 60  
 DB 1 MLLLLAEYLQYKFGVGYLTLRGLSVLTSLSLWLGPMWIRTLQIQIGQAVRND 60  
 QY 61 GPQSHLSKKGTPMGALITLTAISTLLWADLSNRYVWVLLVWTLFGATGWDDYRKV 120  
 DB 61 GPQSHLSKKGTPMGALITLTAISTLLWADLSNRYVWVLLVWTLFGATGWDDYRKV 120  
 QY 121 IEKNSRGLPSRWKRYFQSGVFGIGAAVFLYWTATPIETTLIVPMLKSVEIQIGFFVVL 180  
 DB 121 IEKNSRGLPSRWKRYFQSGVFGIGAAVFLYWTATPIETTLIVPMLKSVEIQIGFFVVL 180  
 QY 181 YFVIVGSSNAVNLTDLGLAIMPVWAGALIGFCYLSGNVKEAYLLIPNVPAGELI 240

DB 181 YFVIVGSSNAVNLTDLGLAIMPVWAGALIGFCYLSGNVKEAYLLIPNVPAGELI 240  
 QY 241 VFCAALVAGLGLFELWNTYPAQVFMGDVGAALGALGTAIVIRQEIYVLFMGVYVME 300  
 DB 241 VFCAALVAGLGLFELWNTYPAQVFMGDVGAALGALGTAIVIRQEIYVLFMGVYVME 300  
 QY 301 TLSVMIQVASFKLTGRVFRMAPIHHPFKGPPDRVIRFWIITVILVLGLATLKL 360  
 DB 301 TLSVMIQVASFKLTGRVFRMAPIHHPFKGPPDRVIRFWIITVILVLGLATLKL 360  
 RESULT 4  
 ABU39977  
 ID ABU39977 standard; protein; 360 AA.  
 XX AC ABU39977;  
 XX DT 19-JUN-2003 (first entry)  
 XX Protein encoded by Prokaryotic essential gene #25504.  
 XX Antisense; prokaryotic essential gene; cell proliferation; drug design.  
 XX *Pseudomonas putida*.  
 XX WO200277183-A2.  
 XX 03-OCT-2002.  
 XX 21-MAR-2002; 2002WO-US009107.  
 XX 21-MAR-2001; 2001US-00815242.  
 XX 06-SEP-2001; 2001US-00948993.  
 XX 25-OCT-2001; 2001US-0342923P.  
 XX 08-FEB-2002; 2002US-00072851.  
 XX 06-MAR-2002; 2002US-0362699P.  
 XX (ELIT-) ELITRA PHARM INC.  
 XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;  
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;  
 XX WPI; 2003-029926/02.  
 DR N-PSDB; ACA43847.  
 XX New antisense nucleic acids, useful for identifying proteins or screening  
 PT for homologous nucleic acids required for cellular proliferation to  
 PT isolate candidate molecules for rational drug discovery programs.  
 XX Claim 25; SEQ ID NO 67901; 1766pp; English.  
 XX The invention relates to an isolated nucleic acid comprising any one of  
 CC the 6213 antisense sequences given in the specification where expression  
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:  
 CC (1) a vector comprising a promoter operably linked to the nucleic acid  
 CC encoding a polypeptide whose expression is inhibited by the antisense  
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated  
 CC polypeptide or its fragment whose expression is inhibited by the  
 CC antisense nucleic acid; (4) an antibody capable of specifically binding  
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular  
 CC proliferation or the activity of a gene in an operon required for  
 CC proliferation; (7) identifying a compound that influences the activity of  
 CC the gene product or that has an activity against a biological pathway  
 CC required for proliferation, or that inhibits cellular proliferation; (8)  
 CC identifying a gene required for cellular proliferation or the biological  
 CC pathway in which a proliferation-required gene or its gene product lies  
 CC or a gene on which the test compound that inhibits proliferation of an  
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a  
 CC compound's activity; (11) a culture comprising strains in which the gene  
 CC product is overexpressed or underexpressed; (12) determining the extent  
 CC to which each of the strains is present in a culture or collection of  
 CC strains; or (13) identifying the target of a compound that inhibits the  
 CC proliferation of an organism. The antisense nucleic acids are useful for  
 CC identifying proteins or screening for homologous nucleic acids required  
 CC for cellular proliferation to isolate candidate molecules for rational  
 CC drug discovery programs, or for screening homologous nucleic acids  
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,  
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of  
 CC the target prokaryotic essential genes. Note: The sequence data for this  
 CC patent did not form part of the printed specification, but was obtained  
 CC in electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 XX SQ Sequence 360 AA;  
 Query Match 99.3%; Score 1824; DB 6; Length 360;  
 Best Local Similarity 99.4%; Pred. No. 6.1e-192;  
 Matches 358; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 MLLLLAEYLQYKFGVGYLTLRGLSVLTSLSLWLGPMWIRTLQIQIGQAVRND 60  
 DB 1 MLLLLAEYLQYKFGVGYLTLRGLSVLTSLSLWLGPMWIRTLQIQIGQAVRND 60  
 QY 61 GPQSHLSKKGTPMGALITLTAISTLLWADLSNRYVWVLLVWTLFGATGWDDYRKV 120  
 DB 61 GPQSHLSKKGTPMGALITLTAISTLLWADLSNRYVWVLLVWTLFGATGWDDYRKV 120  
 QY 121 IEKNSRGLPSRWKRYFQSGVFGIGAAVFLYWTATPIETTLIVPMLKSVEIQIGFFVVL 180  
 DB 121 IEKNSRGLPSRWKRYFQSGVFGIGAAVFLYWTATPIETTLIVPMLKSVEIQIGFFVVL 180  
 QY 181 YFVIVGSSNAVNLTDLGLAIMPVWAGALIGFCYLSGNVKEAYLLIPNVPAGELI 240

CC proliferation of an organism. The antisense nucleic acids are useful for  
CC identifying proteins or screening for homologous nucleic acids required  
CC for cellular proliferation to isolate candidate molecules for rational  
CC drug discovery programs, or for screening homologous nucleic acids  
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,  
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of  
CC the target prokaryotic essential genes. Note: The sequence data for this  
CC patent did not form part of the printed specification, but was obtained  
CC in electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 360 AA;

Query Match 91.9%; Score 1689; DB 6; Length 360;  
Best Local Similarity 90.3%; Pred. No. 4.6e-177; Indels 0; Gaps 0;  
Matches 325; Conservative 19; Mismatches 17;

Qy 1 MLLLAAYLQOQYKGFVFOYLTALSLVLTALSLMLGPMWIRTLQIQGAVRND 60  
Db 1 MLLLAAYLQOQYKGFVFOYLTALSLVLTALSLMLGPMWIRTLQIQGAVRND 60

Qy 61 GPQSHLSKGGTPTMGALILTAITALLWADLSNRYVWVLLFGLAGWDDYRKV 120  
Db 61 GPQSHLSKGGTPTMGALILTAITALLWADLSNRYVWVLLFGLAGWDDYRKV 120

Qy 121 IEKNSRGLPSRWKQYFQVSGVFGAIAVFLYMTAETPIETTLIVPMLKSVRIQLGFFVLT 180  
Db 121 IEKNSRGLPSRWKQYFQVSGVFGAIAVFLYMTAETPIETTLIVPMLKSVRIQLGFFVLT 180

Qy 181 YFVIVGSSNAVNLTDGLDLAIMPVWAGALGIFCYLSGNVFAEYLLIPNPGAGELI 240  
Db 181 YFVIVGSSNAVNLTDGLDLAIMPVWAGALGIFCYLSGNVFAEYLLIPNPGAGELI 240

Qy 241 VFCAALVGAGLGFNFNTYPAQVFMGVDGALGALGTTAVIVROEIVLIFMGVFNME 300  
Db 241 VFCAALVGAGLGFNFNTYPAQVFMGVDGALGALGTTAVIVROEIVLIFMGVFNME 300

Qy 301 TLSVMIOVASFKLTGRVFRMAPIHHPFELKGPDPDRVIVRFTIIVLVLGLATLKL 360  
Db 301 TLSVMIOVASFKLTGRVFRMAPIHHPFELKGPDPDRVIVRFTIIVLVLGLATLKL 360

RESULT 5  
ABU41789  
ID ABU41789 standard; protein; 360 AA.  
XX  
AC ABU41789;  
XX  
DT 19-JUN-2003 (first entry)  
XX  
DE Protein encoded by Prokaryotic essential gene #27316.  
XX  
KW Antisense; prokaryotic essential gene; cell proliferation; drug design.  
XX  
OS Pseudomonas syringae.  
XX  
PN WO200277183-A2.  
XX  
PD 03-OCT-2002.  
XX  
PF 21-MAR-2002; 2002WO-US009107.  
XX  
PR 21-MAR-2001; 2001US-00815242.  
PR 06-SEP-2001; 2001US-00948993.  
PR 25-OCT-2001; 2001US-0342923P.  
PR 08-FEB-2002; 2002US-00072851.  
PR 06-MAR-2002; 2002US-0362699P.  
XX  
PA (ELIT-) ELITRA PHARM INC.  
XX  
PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;  
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu H;

DR WPI; 2003-029926/02.  
DR N-FSDS; ACA45659.  
XX  
PT New antisense nucleic acids, useful for identifying proteins or screening  
PT for homologous nucleic acids required for cellular proliferation to  
PT isolate candidate molecules for rational drug discovery programs.  
XX  
PS Claim 25; SEQ ID NO 69713; 1766pp; English.  
XX  
CC The invention relates to an isolated nucleic acid comprising any one of  
CC the 6213 antisense sequences given in the specification where expression  
CC of the nucleic acid inhibits proliferation of a cell. Also included are:  
CC (1) a vector comprising a promoter operably linked to the nucleic acid  
CC encoding a polypeptide whose expression is inhibited by the antisense  
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated  
CC polypeptide or its fragment whose expression is inhibited by the  
CC antisense nucleic acid; (4) an antibody capable of specifically binding  
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular  
CC proliferation or the activity of a gene in an operon required for  
CC proliferation; (7) identifying a compound that influences the activity of  
CC the gene product or that has an activity against a biological pathway  
CC required for proliferation, or that inhibits cellular proliferation; (8)  
CC identifying a gene required for cellular proliferation or the biological  
CC pathway in which a proliferation-required gene or its gene product lies  
CC or a gene on which the test compound that inhibits proliferation of an  
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a  
CC product's activity; (11) a culture comprising strains in which the gene  
CC product is overexpressed or underexpressed; (12) determining the extent  
CC to which each of the strains is present in a culture or collection of  
CC strains; or (13) identifying the target of a compound that inhibits the  
CC proliferation of an organism. The antisense nucleic acids are useful for  
CC identifying proteins or screening for homologous nucleic acids required  
CC for cellular proliferation to isolate candidate molecules for rational  
CC drug discovery programs, or for screening homologous nucleic acids  
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,  
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of  
CC the target prokaryotic essential genes. Note: The sequence data for this  
CC patent did not form part of the printed specification, but was obtained  
CC in electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 360 AA;

Query Match 90.3%; Score 1658; DB 6; Length 360;  
Best Local Similarity 89.7%; Pred. No. 1.2e-173;  
Matches 323; Conservative 17; Mismatches 20; Indels 0; Gaps 0;

Qy 1 MLLLAAYLQOQYKGFVFOYLTALSLVLTALSLMLGPMWIRTLQIQGAVRND 60  
Db 1 MLLLAAYLQOQYKGFVFOYLTALSLVLTALSLMLGPMWIRTLQIQGAVRND 60

Qy 61 GPQSHLSKGGTPTMGALILTAITALLWADLSNRYVWVLLFGLAGWDDYRKV 120  
Db 61 GPQSHLSKGGTPTMGALILTAITALLWADLSNRYVWVLLFGLAGWDDYRKV 120

Qy 121 IEKNSRGLPSRWKQYFQVSGVFGAIAVFLYMTAETPIETTLIVPMLKSVRIQLGFFVLT 180  
Db 121 IEKNSRGLPSRWKQYFQVSGVFGAIAVFLYMTAETPIETTLIVPMLKSVRIQLGFFVLT 180

Qy 181 YFVIVGSSNAVNLTDGLDLAIMPVWAGALGIFCYLSGNVFAEYLLIPNPGAGELI 240  
Db 181 YFVIVGSSNAVNLTDGLDLAIMPVWAGALGIFCYLSGNVFAEYLLIPNPGAGELI 240

Qy 241 VFCAALVGAGLGFNFNTYPAQVFMGVDGALGALGTTAVIVROEIVLIFMGVFNME 300  
Db 241 VFCAALVGAGLGFNFNTYPAQVFMGVDGALGALGTTAVIVROEIVLIFMGVFNME 300

Qy 301 TLSVMIOVASFKLTGRVFRMAPIHHPFELKGPDPDRVIVRFTIIVLVLGLATLKL 360  
Db 301 TLSVMIOVASFKLTGRVFRMAPIHHPFELKGPDPDRVIVRFTIIVLVLGLATLKL 360

ABU49535  
ID ABU49535 standard; protein; 360 AA.  
XX AC ABU49535;  
XX DT 19-JUN-2003 (first entry)  
XX DE Protein encoded by Prokaryotic essential gene #35062.  
XX KW Antisense; prokaryotic essential gene; cell proliferation; drug design.  
XX OS Vibrio cholerae.  
XX PN WO200277183-A2.  
XX PD 03-OCT-2002.  
XX PF 21-MAR-2002; 2002WO-US009107.  
XX PR 21-MAR-2001; 2001US-00815242.  
XX PR 06-SEP-2001; 2001US-00948993.  
XX PR 25-OCT-2001; 2001US-0342923P.  
XX PR 08-FEB-2002; 2002US-00072851.  
XX PR 06-MAR-2002; 2002US-0362693P.  
XX PA (ELIT-) ELITRA PHARM INC.  
XX PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;  
XX PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;  
XX DR WPI: 2003-029926/02.  
XX DR N-PSDB; ACN53405.  
XX PT New antisense nucleic acids, useful for identifying proteins or screening  
XX PT for homologous nucleic acids required for cellular proliferation to  
XX PT isolate candidate molecules for rational drug discovery programs.  
XX PS Claim 25; SEQ ID NO 77459; 1766pp; English.  
XX CC The invention relates to an isolated nucleic acid comprising any one of  
XX CC the 6213 antisense sequences given in the specification where expression  
XX CC of the nucleic acid inhibits proliferation of a cell. Also included are:  
XX CC (1) a vector comprising a promoter operably linked to the nucleic acid  
XX CC encoding a polypeptide whose expression is inhibited by the antisense  
XX CC nucleic acid; (2) a host cell containing the vector; (3) an isolated  
XX CC polypeptide or its fragment whose expression is inhibited by the  
XX CC antisense nucleic acid; (4) an antibody capable of specifically binding  
XX CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular  
XX CC proliferation or the activity of a gene in an operon required for  
XX CC proliferation; (7) identifying a compound that influences the activity of  
XX CC the gene product or that has an activity against a biological pathway  
XX CC required for proliferation, or that inhibits cellular proliferation; (8)  
XX CC identifying a gene required for cellular proliferation or the biological  
XX CC pathway in which a proliferation-required gene or its gene product lies  
XX CC or a gene on which the test compound that inhibits proliferation of an  
XX CC organism acts; (9) manufacturing an antibiotic; (10) profiling a  
XX CC compound's activity; (11) a culture comprising strains in which the gene  
XX CC product is overexpressed or underexpressed; (12) determining the extent  
XX CC to which each of the strains is present in a culture or collection of  
XX CC strains; or (13) identifying the target of a compound that inhibits the  
XX CC proliferation of an organism. The antisense nucleic acids are useful for  
XX CC identifying proteins or screening for homologous nucleic acids required  
XX CC for cellular proliferation to isolate candidate molecules for rational  
XX CC drug discovery programs, or for screening homologous nucleic acids  
XX CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,  
XX CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of  
XX CC the target prokaryotic essential genes. Note: the sequence data for this  
XX CC patent did not form part of the printed specification, but was obtained  
XX CC in electronic format directly from WIPO at  
XX CC ftp.wipo.int/pub/published\_pct\_sequences  
XX CC Sequence 360 AA;

Query Match 70.3%; Score 1292; DB 6; Length 360;  
Best Local Similarity 66.9%; Pred. No. 2.6e-133;  
Matches 241; Conservative 51; Mismatches 68; Indels 0; Gaps 0;  
QY 1 MLLLAELYQQFYKGFVQYLTIRILSVLTALSLSLWLGPMIRTLQIPQIGQAVRND 60  
DB 1 MIITLAEELQPVFSPFELFEYLSFRAIVSILTALGSLWGMSPRMKRLQMLQIGQVVRNE 60  
QY 61 GPQSHLSKSGTPTMGALILTAISTLWADLSNRYVWVLTLLFGAIGWVDYRKV 120  
DB 61 GPESHLSKSGTPTMGGVWILAAITITVLLWADLTNPYVWAVLAVLLGYGAVGFVDYRKV 120  
QY 121 IEKNSRGLPSRWKYFWQSVFGIGAAVFLVNTAETPIETTLIVPMLKSVIQLGIFVVL 180  
DB 121 VRKNTDGLIARWKYFWQSAIALVAPALYAHGQDTAATQLVVFFFDVNVFQLGLMIVLT 180  
QY 181 YFVIVGSSNAVNLTDGLGLAIMPVTWVAGALGIFCLYLSGNVKFAEYLLIPNVPGAGELI 240  
DB 181 YFVIVGSSNAVNLTDGLGLAIMPVTVVAAGFAAIAWAIGNVNPANYLHIPYIPHSSELV 240  
QY 241 VFCAALVAGALGFLWNTYPAQVFMGDVGALGALGALGTIAVIVRQEI VLFIMGGVFNME 300  
DB 241 VVCTAMVAGALGFLWNTYPAQVFMGDVGALGALGALGTIAVLRQEFVLMGGVFNME 300  
QY 301 TLSVMTQVASFKLITGRVFRMAPIHFFELKGFDPDVIVRFWIITVILVILGLATLKL 360  
DB 301 TLSVILQVSGYKLRGQRIIFRMAPIHFFELKGFDPDVIVRFWIISIVLVILGLATLKL 360  
RESULT 7  
ABM68871  
ID ABM68871 standard; protein; 361 AA.  
XX AC ABM68871;  
XX DT 20-NOV-2003 (first entry)  
XX DE Photorhabdus luminescens protein sequence #1968.  
XX KW Antibacterial; fungicide; insecticide; polymorphism; genetic analysis;  
XX KW detection; food; gene expression; plant; animal; microorganism; toxin;  
XX KW antibiotic; biopesticide; virulence factor; disease model; plague;  
XX KW whooping cough.  
XX OS Photorhabdus luminescens.  
XX PN WO200294867-A2.  
XX PD 28-NOV-2002.  
XX PF 07-FEB-2002; 2002WO-IB003040.  
XX PR 07-FEB-2001; 2001FR-00001659.  
XX PA (INSP ) INST PASTEUR.  
XX PA (CNRS ) CNRS CENT NAT RECH SCI.  
XX PI Duchaud E, Taourit S, Glaser P, Frangeul L, Kunst F, Danchin A;  
XX PI Buchrieser C;  
XX WPI: 2003-148459/14.  
XX DR Genomic sequence of Photorhabdus luminescens and encoded polypeptides,  
XX PT useful e.g. as therapeutic antimicrobials and agricultural pesticides.  
XX PS Claim 2; SEQ ID NO 1968; 1205pp; French.  
XX CC The invention relates to the isolation of genes and their encoded  
XX CC proteins from Photorhabdus luminescens. The isolated sequences are  
XX CC sources of probes and primers for detecting the genome of *P. luminescens*  
XX CC and related species; to study polymorphisms; for gene analysis and for  
XX CC detection/amplification of the genes. Antibodies (Ab) raised against the  
XX CC polypeptides encoded by the genes are used for detection/identification

CC of P. luminescens, e.g. in foods. The genes, proteins, Ab and cells that  
CC carry a gene-containing vector are used to select compounds that  
CC modulate, regulate, induce or inhibit expression of the genes in plants,  
CC animals or microorganisms other than P. luminescens and are able to alter  
CC response or sensitivity to toxins and antibiotics produced by P.  
CC luminescens. Cells transformed to express the genes are useful for  
CC recombinant production of the proteins, particularly toxins and  
CC antibacterials useful as insecticides, bactericides and fungicides. The  
CC genes, proteins, vectors containing the genes and Ab are also useful  
CC therapeutically to treat microbial infection by bacteria or fungi that  
CC are sensitive to P. luminescens-encoded toxins or antibiotics and as  
CC biopesticides. Other uses of the genes and the proteins are as virulence  
CC factors and for identifying targets of human diseases for which P.  
CC luminescens is a model (particularly plague and whooping cough). This  
CC sequence represents one of the isolated P. luminescens proteins  
XX  
XX  
SQ Sequence 361 AA;

Query Match 69.7%; Score 1280; DB 6; Length 361;  
Best Local Similarity 66.4%; Pred. No. 5.5e-132;  
Matches 239; Conservative 55; Mismatches 66; Indels 0; Gaps 0;

QY 1 MLLLLAEVLQOYKGFVFOYLTIRGLSVLTALSLWLGPMWIRTLQIPQIGQVRND 60  
DB 1 MLVWLABYLVKXHSFNVFSLTFRAIVSLTALALWNGPRMTAFLOKLQIGQVRND 60  
QY 61 GPQSHLSKKGTPMTGGALILTAIAISTLLWADLSNRYVWVVLVTLFGAIGWDDYRKV 120  
DB 61 GPESHFSKRGTPMTGGVILFAITVSTLLWANLANPVWFSLVLLGVGAIGFVDDYRKI 120  
QY 121 IEKNSRGLPSRWKFWQSVFGIAGAFVLYMTAETPIETTLIVPMLKSVETQLGFVVLT 180  
DB 121 VRKDTKGLIARWKYFQWSVGLALGVAFMYSGFKDTPATQLWVPFFKXVMPQLGVLYLT 180  
QY 181 YFVIVGSSNAVNLTDGLDGLAIMPTVMVAGALGIFCVLSGNVKFAFYLIPNVPGAGELI 240  
DB 181 YFVIVGTSNAVNLTDGLDGLAIMPTVFVAGFALVAVATGNVFNALYHPIYLSHAGELV 240  
QY 241 VFCAALVAGLGLFWNTYPAQVFMGVDGVALGALGTTIAVIRQEVILFMGQVFNVE 300  
DB 241 IVCTAIVAGLGLFWNTYPAQVFMGVDGVALGALGTTIAVLRQEVILFMGQVFNVE 300  
QY 301 TLSVMIVQVAFKLTGRVFRMAPTHHHFELKGPDPVIVRFVWIIIVLVLGLATLKL 360  
DB 301 TLSVILQVGFKLGRQIFRMAPTHHHFELKGPDPVIVRFVWIIIVLVLGLATLKL 360

RESULT 8  
ABU38972  
ID ABU38972 standard; protein; 360 AA.  
XX  
XX AC ABU38972;  
XX  
XX DT 19-JUN-2003 (first entry)  
XX  
XX DE Protein encoded by prokaryotic essential gene #24499.  
XX  
XX KW Antisense; prokaryotic essential gene; cell proliferation; drug design.  
XX  
XX OS Pasteurella multocida.  
XX  
XX PN WO200277183-A2.  
XX  
XX PD 03-OCT-2002.  
XX  
XX PF 21-MAR-2002; 2002WO-US009107.  
XX  
XX PR 21-MAR-2001; 2001US-00915242.  
XX  
XX PR 06-SEP-2001; 2001US-00948993.  
XX  
XX PR 25-OCT-2001; 2001US-0342923P.  
XX  
XX PR 08-FEB-2002; 2002US-00072851.  
XX  
XX PR 06-MAR-2002; 2002US-0362699P.

(ELIT-) ELITRA PHARM INC.  
Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;  
Wail D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;  
WPI; 2003-029926/02.  
DR N-PSDB; ACA42842.  
XX  
XX New antisense nucleic acids, useful for identifying proteins or screening  
XX for homologous nucleic acids required for cellular proliferation to  
XX isolate candidate molecules for rational drug discovery programs.  
Claim 25; SEQ ID NO 66896; 1766bp; English.  
CC  
CC The invention relates to an isolated nucleic acid comprising any one of  
CC the 6213 antisense sequences given in the specification where expression  
CC of the nucleic acid inhibits proliferation of a cell. Also included are:  
CC (1) a vector comprising a promoter operably linked to the nucleic acid  
CC encoding a polypeptide whose expression is inhibited by the antisense  
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated  
CC antisense nucleic acid; (4) an antibody capable of specifically binding  
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular  
CC proliferation or the activity of a gene in an operon required for  
CC proliferation; (7) identifying a compound that influences the activity of  
CC the gene product or that has an activity against a biological pathway  
CC required for proliferation, or that inhibits cellular proliferation; (8)  
CC identifying a gene required for cellular proliferation or the biological  
CC pathway in which a proliferation-required gene or its gene product lies  
CC or a gene on which the test compound that inhibits proliferation of an  
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a  
CC compound's activity; (11) a culture comprising strains in which the gene  
CC product is overexpressed or underexpressed; (12) determining the extent  
CC to which each of the strains is present in a culture or collection of  
CC strains; or (13) identifying the target of a compound that inhibits the  
CC proliferation of an organism. The antisense nucleic acids are useful for  
CC identifying proteins or screening for homologous nucleic acids required  
CC for cellular proliferation to isolate candidate molecules for rational  
CC drug discovery programs, or for screening homologous nucleic acids  
CC required for proliferation in cells other than S. aureus, S. typhimurium,  
CC K. pneumoniae or P. aeruginosa. The present sequence is encoded by one of  
CC the target prokaryotic essential genes. Note: The sequence data for this  
CC patent did not form part of the printed specification, but was obtained  
CC in electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences  
XX  
XX SQ Sequence 360 AA;

Query Match 69.4%; Score 1274; DB 6; Length 360;  
Best Local Similarity 65.8%; Pred. No. 2.5e-131;  
Matches 237; Conservative 57; Mismatches 66; Indels 0; Gaps 0;

QY 1 MLLLLAEVLQOYKGFVFOYLTIRGLSVLTALSLWLGPMWIRTLQIPQIGQVRND 60  
DB 1 MLVWLABYLVKXHSFNVFSLTFRAIVSLTALALWNGPRMTAFLOKLQIGQVRND 60  
QY 61 GPQSHLSKKGTPMTGGALILTAIAISTLLWADLSNRYVWVVLVTLFGAIGWDDYRKV 120  
DB 61 GPESHFSKRGTPMTGGVILFAITVSTLLWANLANPVWFSLVLLGVGAIGFVDDYRKI 120  
QY 121 IEKNSRGLPSRWKFWQSVFGIAGAFVLYMTAETPIETTLIVPMLKSVETQLGFVVLT 180  
DB 121 TRKNTDGLIARWKYFQWSVGLALGVAFMYSGFKDTPATQLWVPFFKXVMPQLGVLYLT 180  
QY 181 YFVIVGSSNAVNLTDGLDGLAIMPTVMVAGALGIFCVLSGNVKFAFYLIPNVPGAGELI 240  
DB 181 YFVIVGTSNAVNLTDGLDGLAIMPTVFVAGFALVAVATGNVFNALYHPIYLSHAGELV 240  
QY 241 VFCAALVAGLGLFWNTYPAQVFMGVDGVALGALGTTIAVIRQEVILFMGQVFNVE 300  
DB 241 IVCTAIVAGLGLFWNTYPAQVFMGVDGVALGALGTTIAVLRQEVILFMGQVFNVE 300  
QY 301 TLSVMIVQVAFKLTGRVFRMAPTHHHFELKGPDPVIVRFVWIIIVLVLGLATLKL 360  
DB 301 TLSVILQVGFKLGRQIFRMAPTHHHFELKGPDPVIVRFVWIIIVLVLGLATLKL 360

Query Match 69.4%; Score 1274; DB 6; Length 360;  
Best Local Similarity 65.8%; Pred. No. 2.5e-131;  
Matches 237; Conservative 57; Mismatches 66; Indels 0; Gaps 0;

QY 1 MLLLLAEVLQOYKGFVFOYLTIRGLSVLTALSLWLGPMWIRTLQIPQIGQVRND 60  
DB 1 MLVWLABYLVKXHSFNVFSLTFRAIVSLTALALWNGPRMTAFLOKLQIGQVRND 60  
QY 61 GPQSHLSKKGTPMTGGALILTAIAISTLLWADLSNRYVWVVLVTLFGAIGWDDYRKV 120  
DB 61 GPESHFSKRGTPMTGGVILFAITVSTLLWANLANPVWFSLVLLGVGAIGFVDDYRKI 120  
QY 121 IEKNSRGLPSRWKFWQSVFGIAGAFVLYMTAETPIETTLIVPMLKSVETQLGFVVLT 180  
DB 121 TRKNTDGLIARWKYFQWSVGLALGVAFMYSGFKDTPATQLWVPFFKXVMPQLGVLYLT 180  
QY 181 YFVIVGSSNAVNLTDGLDGLAIMPTVMVAGALGIFCVLSGNVKFAFYLIPNVPGAGELI 240  
DB 181 YFVIVGTSNAVNLTDGLDGLAIMPTVFVAGFALVAVATGNVFNALYHPIYLSHAGELV 240  
QY 241 VFCAALVAGLGLFWNTYPAQVFMGVDGVALGALGTTIAVIRQEVILFMGQVFNVE 300  
DB 241 IVCTAIVAGLGLFWNTYPAQVFMGVDGVALGALGTTIAVLRQEVILFMGQVFNVE 300  
QY 301 TLSVMIVQVAFKLTGRVFRMAPTHHHFELKGPDPVIVRFVWIIIVLVLGLATLKL 360  
DB 301 TLSVILQVGFKLGRQIFRMAPTHHHFELKGPDPVIVRFVWIIIVLVLGLATLKL 360

[illegible]

CC that are specific for particular species of microorganisms can be used as  
CC probes to identify particular microorganism species in clinical  
CC specimens. AAH81295 to AAH81487 encode the *Escherichia coli* proteins  
CC given in AAG98239 to AAG98431, and AAH81488 to AAH81491 represent  
CC oligonucleotides, which are used in the exemplification of the present  
CC invention  
XX  
XX Sequence 360 AA;  
Query Match 69.0%; Score 1268; DB 4; Length 360;  
Best Local Similarity 66.4%; Pred. No. 1.2e-130;  
Matches 239; Conservative 51; Mismatches 70; Indels 0; Gaps 0;  
Qy 1 MLLLAELVLCQFVKGFGVFOYLTGRLTSLTSLSLGPMWIRTLQIPQIGQAVRND 60  
Db 1 MLVWLAHLVKYSGFVFSYLTFRALVSLTALFISLWMPRIAHQLKLSFGQVRND 60  
Qy 61 GPQSHLSKGGTPTMGGALITATAISLTLWADLSNRYVWVLLFGLAGWDDYRKV 120  
Db 61 GPESHFSKGGTPTMGGIMLTATVISVLLWAYSPNPMVLCVLLVGVGIGFVDDYRKV 120  
Qy 121 IEKNSRGLPSRWKYSQVFGIGAAVFLYMTAETPIETTLIVPMLKSVETQLGIEFVVL 180  
Db 121 VRKDTKGLIARWKYFWMVSVIALGVAFALYLAGKDTPTATQLVVPFFKDVMPQLGLFYLLA 180  
Qy 181 YFVIVGSSNAVNLTDGLDLAIMPTVMVAGALGIFCYLSGNVRFAEYLLIPNVPAGELI 240  
Db 181 YFVIVGTGNAVNLTDGLDLAIMPTVFAGGFALVAVATGNMNFASYLHIPYLRHAGELV 240  
Qy 241 VFCAALVAGLGFWMFYTPAQVFMGVDGVALALGAALGTATVIVRQEVILFIMGVFWME 300  
Db 241 IVCTAIVAGLGLFWNTYTPAQVFMGVDGVALGALGIIATVLLRQBFLLVIMGVFWVE 300  
Qy 301 TLSVMIQVASFKLTRGRVFRMAPIHHPHFELKGPDPDRVIVRFWIITVILVLIGLATLKL 360  
Db 301 TLSVILQVGSFKLRGQRIFRMAPIHHPHFELKGPDPDRVIVRFWIISLMLVLIGLATLKV 360  
RESULT 11  
ABU28488  
ID ABU28488 standard; protein; 360 AA.  
AC ABU28488;  
XX  
XX 19-JUN-2003 (first entry)  
DE Protein encoded by Prokaryotic essential gene #14015.  
XX Antisense; prokaryotic essential gene; cell proliferation; drug design.  
XX *Escherichia coli*.  
XX WO200277183-A2.  
XX 03-OCT-2002.  
XX 21-MAR-2002; 2002WO-US009107.  
XX 21-MAR-2001; 2001US-00815242.  
PR 06-SEP-2001; 2001US-00948993.  
PR 25-OCT-2001; 2001US-0342923P.  
PR 08-FEB-2002; 2002US-00072851.  
PR 06-MAR-2002; 2002US-0362699P.  
XX (ELIT-) ELITRA PHARM INC.  
XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;  
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;  
XX WPI; 2003-029926/02.  
DR N-PSDB; ACA32358.  
XX  
XX New antisense nucleic acids, useful for identifying proteins or screening

PT for homologous nucleic acids required for cellular proliferation to  
PT isolate candidate molecules for rational drug discovery programs.  
XX  
XX Claim 25; SEQ ID NO 56412; 1766pp; English.  
CC  
CC The invention relates to an isolated nucleic acid comprising any one of  
CC the 6213 antisense sequences given in the specification where expression  
CC of the nucleic acid inhibits proliferation of a cell. Also included are:  
CC (1) a vector comprising a promoter operably linked to the nucleic acid  
CC encoding a polypeptide whose expression is inhibited by the antisense  
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated  
CC polypeptide or its fragment whose expression is inhibited by the  
CC antisense nucleic acid; (4) an antibody capable of specifically binding  
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular  
CC proliferation or the activity of a gene in an operon required for  
CC proliferation; (7) identifying a compound that influences the activity of  
CC the gene product or that has an activity against a biological pathway  
CC required for proliferation, or that inhibits cellular proliferation; (8)  
CC identifying a gene required for cellular proliferation or the biological  
CC pathway in which a proliferation-required gene or its gene product lies  
CC or a gene on which the test compound that inhibits proliferation of an  
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a  
CC compound's activity; (11) a culture comprising strains in which the gene  
CC product is overexpressed or underexpressed; (12) determining the extent  
CC to which each of the strains is present in a culture or collection of  
CC strains; or (13) identifying the target of a compound that inhibits the  
CC proliferation of an organism. The antisense nucleic acids are useful for  
CC identifying proteins or screening for homologous nucleic acids required  
CC for cellular proliferation to isolate candidate molecules for rational  
CC drug discovery programs, or for screening homologous nucleic acids  
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,  
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of  
CC the target prokaryotic essential genes. Note: the sequence data for this  
CC patent did not form part of the printed specification, but was obtained  
CC in electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences  
XX  
XX Sequence 360 AA;  
Query Match 69.0%; Score 1268; DB 6; Length 360;  
Best Local Similarity 66.4%; Pred. No. 1.2e-130;  
Matches 239; Conservative 51; Mismatches 70; Indels 0; Gaps 0;  
Qy 1 MLLLAELVLCQFVKGFGVFOYLTGRLTSLTSLTALSLSLWLPMMIRTLQIPQIGQAVRND 60  
Db 1 MLVWLAHLVKYSGFVFSYLTFRALVSLTALFISLWMPRIAHQLKLSFGQVRND 60  
Qy 61 GPQSHLSKGGTPTMGGALITATAISLTLWADLSNRYVWVLLFGLAGWDDYRKV 120  
Db 61 GPESHFSKGGTPTMGGIMLTATVISVLLWAYSPNPMVLCVLLVGVGIGFVDDYRKV 120  
Qy 121 IEKNSRGLPSRWKYSQVFGIGAAVFLYMTAETPIETTLIVPMLKSVETQLGIEFVVL 180  
Db 121 VRKDTKGLIARWKYFWMVSVIALGVAFALYLAGKDTPTATQLVVPFFKDVMPQLGLFYLLA 180  
Qy 181 YFVIVGSSNAVNLTDGLDLAIMPTVMVAGALGIFCYLSGNVRFAEYLLIPNVPAGELI 240  
Db 181 YFVIVGTGNAVNLTDGLDLAIMPTVFAGGFALVAVATGNMNFASYLHIPYLRHAGELV 240  
Qy 241 VFCAALVAGLGLFWNTYTPAQVFMGVDGVALALGAALGTATVIVRQEVILFIMGVFWVE 300  
Db 241 IVCTAIVAGLGLFWNTYTPAQVFMGVDGVALGALGIIATVLLRQBFLLVIMGVFWVE 300  
Qy 301 TLSVMIQVASFKLTRGRVFRMAPIHHPHFELKGPDPDRVIVRFWIITVILVLIGLATLKL 360  
Db 301 TLSVILQVGSFKLRGQRIFRMAPIHHPHFELKGPDPDRVIVRFWIISLMLVLIGLATLKV 360  
RESULT 12  
ABU49996  
ID ABU49996 standard; protein; 360 AA.  
XX  
XX AC ABU49996;

XX 19-JUN-2003 (first entry)  
XX Protein encoded by Prokaryotic essential gene #35523.  
XX Antisense; prokaryotic essential gene; cell proliferation; drug design.  
XX Yersinia pestis.  
OS WO200277183-A2.  
PN 03-OCT-2002.  
XX 21-MAR-2002; 2002WO-US009107.  
XX 21-MAR-2001; 2001US-00815242.  
XX 06-SEP-2001; 2001US-00948993.  
XX 25-OCT-2001; 2001US-0342923P.  
XX 08-FEB-2002; 2002US-00072851.  
XX 06-MAR-2002; 2002US-0362699P.  
XX (ELIT-) ELITRA PHARM INC.  
XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;  
XX Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;  
XX WPI; 2003-029926/02.  
XX N-PSDB; ACA53866.  
XX New antisense nucleic acids, useful for identifying proteins or screening  
XX for homologous nucleic acids required for cellular proliferation to  
XX isolate candidate molecules for rational drug discovery programs.  
XX Claim 25; SEQ ID NO 77920; 1766pp; English.  
XX The invention relates to an isolated nucleic acid comprising any one of  
XX the 6213 antisense sequences given in the specification where expression  
XX of the nucleic acid inhibits proliferation of a cell. Also included are:  
XX (1) a vector comprising a promoter operably linked to the nucleic acid  
XX encoding a polypeptide whose expression is inhibited by the antisense  
XX nucleic acid; (2) a host cell containing the vector; (3) an isolated  
XX antisense nucleic acid; (4) an antibody capable of specifically binding  
XX the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular  
XX proliferation or the activity of a gene in an operon required for  
XX proliferation; (7) identifying a compound that influences the activity of  
XX the gene product or that has an activity against a biological pathway  
XX required for proliferation, or that inhibits cellular proliferation; (8)  
XX identifying a gene required for cellular proliferation or the biological  
XX pathway in which a proliferation-required gene or its gene product lies  
XX or a gene on which the test compound that inhibits proliferation of an  
XX organism acts; (9) manufacturing an antibiotic; (10) profiling a  
XX compound's activity; (11) a culture comprising strains in which the gene  
XX product is overexpressed or underexpressed; (12) determining the extent  
XX to which each of the strains is present in a culture or collection of  
XX strains; or (13) identifying the target of a compound that inhibits the  
XX proliferation of an organism. The antisense nucleic acids are useful for  
XX identifying proteins or screening for homologous nucleic acids required  
XX for cellular proliferation to isolate candidate molecules for rational  
XX drug discovery programs, or for screening homologous nucleic acids  
XX required for proliferation in cells other than *S. aureus*, *S. typhimurium*,  
XX *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of  
XX the target prokaryotic essential genes. Note: The sequence data for this  
XX patent did not form part of the printed specification, but was obtained  
XX in electronic format directly from WIPO at  
XX ftp.wipo.int/pub/published\_pct\_sequences  
XX Sequence 360 AA;

Query Match 68.9%; Score 1265; DB 6; Length 360;  
Best Local Similarity 65.6%; Pred. No. 2,5e-130;  
Matches 236; Conservative 54; Mismatches 70; Indels 0; Gaps 0;

QY 1 MLLLLAEYLQOFYKGVFOYLTLRGILSVLTALSLWLGPMWIRTQIPQIGQAVRND 60  
DB 1 MLVWLAEVLKPYSGFNVFSYLTFRATVSLTALPISLWMGPHLAWLQKLOIGQVVRND 60  
QY 61 GPQSHLSKSGTMTGALITLTAISTLWADLSNRYVNVVLTLLFGAIGVDDYRKV 120  
DB 61 GPESHFSKSGTMTGALITLTAISTLWADLSNRYVNVVLTLLFGAIGVDDYRKV 120  
QY 121 IEKNSRGLPSRWKYFQWQVFGIGAAVFLYMTAETPIETTLIVPMLKSVBIQLGIFPVLT 180  
DB 121 VRKNTKGLIARWKYFQWQVFGIGAAVFLYMTAETPIETTLIVPMLKSVBIQLGIFPVLT 180  
QY 181 YFVIVGSSNAVNLTDGLDLAIMPVTVAGALGIFCYLSGNVYKFAEYLIPNVPAGELI 240  
DB 181 YFVIVGSSNAVNLTDGLDLAIMPVTVAGALGIFCYLSGNVYKFAEYLIPNVPAGELI 240  
QY 241 VFCAALVAGLGLFMTNTPVPAQVFMGDVGLAALGALGTIAIVIQEIVLFMGVGVVME 300  
DB 241 IVCTAIVAGLGLFMTNTPVPAQVFMGDVGLAALGALGTIAIVIQEIVLFMGVGVVME 300  
QY 301 TLSVMIQVASFKLITGRVFRMAPIHFFHFKWPPRIVRFRWIIITVILVLGLATLKL 360  
DB 301 TLSVMIQVASFKLITGRVFRMAPIHFFHFKWPPRIVRFRWIIITVILVLGLATLKL 360  
RESULT 13  
ABU48094  
ID ABU48094 standard; protein; 360 AA.  
XX AC ABU48094;  
XX XX  
XX DT 19-JUN-2003 (first entry)  
XX DE Protein encoded by Prokaryotic essential gene #33621.  
XX KW Antisense; prokaryotic essential gene; cell proliferation; drug design.  
XX OS Salmomella typhi.  
XX PN WO200277183-A2.  
XX PD 03-OCT-2002.  
XX PF 21-MAR-2002; 2002WO-US009107.  
XX PR 21-MAR-2001; 2001US-00815242.  
XX PR 06-SEP-2001; 2001US-00948993.  
XX PR 25-OCT-2001; 2001US-0342923P.  
XX PR 08-FEB-2002; 2002US-00072851.  
XX PR 06-MAR-2002; 2002US-0362699P.  
XX PA (ELIT-) ELITRA PHARM INC.  
XX PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;  
XX Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;  
XX XX WPI; 2003-029926/02.  
XX DR N-PSDB; ACA51964.  
XX PT New antisense nucleic acids, useful for identifying proteins or screening  
XX for homologous nucleic acids required for cellular proliferation to  
XX isolate candidate molecules for rational drug discovery programs.  
XX PS Claim 25; SEQ ID NO 76018; 1766pp; English.  
XX CC The invention relates to an isolated nucleic acid comprising any one of  
XX the 6213 antisense sequences given in the specification where expression  
XX of the nucleic acid inhibits proliferation of a cell. Also included are:  
XX (1) a vector comprising a promoter operably linked to the nucleic acid  
XX encoding a polypeptide whose expression is inhibited by the antisense  
XX nucleic acid; (2) a host cell containing the vector; (3) an isolated  
XX polypeptide or its fragment whose expression is inhibited by the  
XX antisense nucleic acid; (4) an antibody capable of specifically binding  
XX the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular  
XX proliferation or the activity of a gene in an operon required for  
XX proliferation; (7) identifying a compound that influences the activity of  
XX the gene product or that has an activity against a biological pathway  
XX required for proliferation, or that inhibits cellular proliferation; (8)  
XX identifying a gene required for cellular proliferation or the biological  
XX pathway in which a proliferation-required gene or its gene product lies  
XX or a gene on which the test compound that inhibits proliferation of an  
XX organism acts; (9) manufacturing an antibiotic; (10) profiling a  
XX compound's activity; (11) a culture comprising strains in which the gene  
XX product is overexpressed or underexpressed; (12) determining the extent  
XX to which each of the strains is present in a culture or collection of  
XX strains; or (13) identifying the target of a compound that inhibits the  
XX proliferation of an organism. The antisense nucleic acids are useful for  
XX identifying proteins or screening for homologous nucleic acids required  
XX for cellular proliferation to isolate candidate molecules for rational  
XX drug discovery programs, or for screening homologous nucleic acids  
XX required for proliferation in cells other than *S. aureus*, *S. typhimurium*,  
XX *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of  
XX the target prokaryotic essential genes. Note: The sequence data for this  
XX patent did not form part of the printed specification, but was obtained  
XX in electronic format directly from WIPO at  
XX ftp.wipo.int/pub/published\_pct\_sequences  
XX Sequence 360 AA;



Db 181 YFVIVGTGNVNLTDGLDLAIMPVFAAGFALVAMATGNMFANYLHPIYLRHAGELV 240  
 QY 241 VFCALVAGAGLGFNFNTYPAQVFMGDVGALGALGTIAIVVROEIVLFGVFMVE 300  
 Db 241 IVCTAIVGAGLGFNFNTYPAQVFMGDVGSLGALGALGIIAVLRQEFLLVIMGGVFVE 300  
 QY 301 TLSVNIQVASFKLTKGRVFRKAPIHHFELKGPDPVIVRFWIITVILVIGLATLKL 360  
 Db 301 TLSVILQVGSFKLRQRIFRMAPIHHHYELKGPDPVIVRFWIISLMLVIGLATLKL 360

RESULT 15  
 AAU38482  
 ID AAU38482 standard; protein; 360 AA.  
 AC AAU38482;  
 DT 14-FEB-2002 (first entry)  
 XX Salmonella typhi cellular proliferation protein #373.  
 DE Antisense; prokaryotic cellular proliferation protein; antibiotic;  
 KW antibacterial; drug design.  
 XX Salmonella typhi.  
 OS WO200170955-A2.  
 PN 27-SEP-2001.  
 PD 21-MAR-2001; 2001WO-US009180.  
 PF 21-MAR-2000; 2000US-0191078P.  
 PR 23-MAY-2000; 2000US-0206848P.  
 PR 26-MAY-2000; 2000US-0207727P.  
 PR 23-OCT-2000; 2000US-0242578P.  
 PR 27-NOV-2000; 2000US-0253625P.  
 PR 22-DEC-2000; 2000US-0257931P.  
 PR 16-FEB-2001; 2001US-0269308P.  
 XX (ELIT-) ELITRA PHARM INC.  
 PA Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;  
 PI Yamamoto RT, Xu HH;  
 PI WPI; 2001-611495/70.  
 DR N-PSDB; AAS56341.  
 XX New polynucleotides for the identification and development of  
 PT antibiotics, comprise sequences of antisense nucleic acids.  
 XX Example 3; SEQ ID NO 14075; 51pp; English.

Seq Sequence 360 AA;  
 Query Match 68.5%; Score 1258; DB 4; Length 360;  
 Best Local Similarity 65.8%; Pred. No. 1.5e-129;  
 Matches 237; Conservative 52; Mismatches 71; Indels 0; Gaps 0;  
 QY 1 MLLLLAEYLOQFYKGFVFOYLTGRILSVLTALSLSLMLGFWMTLQIPQIGQAVRND 60  
 Db 1 MLVLAELHLVKYSGFNVSFLTFRATVLSLTALFISLWGPMTARLQKLSFGQVVRND 60  
 QY 61 GPQSHLSKKGTPMGGALILTAISTLLWADLSNRYVWVLLVTLFPGAIGWVDDYRKV 120  
 Db 61 GPESHFSKKGTPMGGIMILTAIVISVLLWVPSNPYVWCVLVVLLIGYGIIGFVDDYRKV 120  
 QY 121 IEKNSRGLPSRMKYFWQSVFGGAAVFLVMTAETPIETTLIVPMLKSVIQLGIFVVL 180  
 Db 121 VRKDTKGLIARWKYFWMSVIALGVAFALYLVGKDTPTATQLVVPFFKDVMPOLGLFYLLS 180  
 QY 181 YFVIVGSSNAVNLTDGLDLAIMPVFAAGFALVAMATGNMFANYLHPIYLRHAGELV 240  
 Db 181 YFVIVGTGNVNLTDGLDLAIMPVFAAGFALVAMATGNMFANYLHPIYLRHAGELV 240  
 QY 241 VFCALVAGAGLGFNFNTYPAQVFMGDVGALGALGTIAIVVROEIVLFGVFMVE 300  
 Db 241 IVCTAIVGAGLGFNFNTYPAQVFMGDVGSALGALGIIAVLRQEFLLVIMGGVFVE 300  
 QY 301 TLSVNIQVASFKLTKGRVFRMAPIHHHFEELKGPDPVIVRFWIITVILVIGLATLKL 360  
 Db 301 TLSVILQVGSFKLRQRIFRMAPIHHHYELKGPDPVIVRFWIISLMLVIGLATLKL 360

Search completed: May 7, 2004, 08:39:26  
 Job time : 63 secs

The invention relates to antisense inhibitors of genes essential to prokaryotic cellular proliferation, their use in identifying the genes, their use in the discovery of novel antibiotics, the essential genes themselves and the encoded proteins. The prokaryotes used are Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The invention is also useful for the identification of potential new targets for antibiotic development. The antisense nucleic acids can also be used to identify proteins used in proliferation, to express these proteins, and to obtain antibodies capable of binding to the expressed proteins. The proteins can be used to screen compounds in rational drug discovery programmes. The antisense nucleic acid sequence is also useful to screen for homologous nucleic acids which are required for cell proliferation in a wide variety of organisms. The present sequence represents an essential prokaryotic cellular proliferation protein. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 9, 2004, 13:18:04 ; Search time 3219 Seconds  
(without alignments)  
10046.829 Million cell updates/sec

Title: US-10-089-787-1  
Perfect score: 1083  
Sequence: 1 atgtctgtgtgtgcgcga.....ccaccttgaagctgcttga 1083

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues  
Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST:  
1: em\_estba.\*  
2: em\_esthum.\*  
3: em\_estin.\*  
4: em\_estmu.\*  
5: em\_estov.\*  
6: em\_estpl.\*  
7: em\_estro.\*  
8: em\_hic.\*  
9: gb\_est1.\*  
10: gb\_est2.\*  
11: gb\_hic.\*  
12: gb\_est3.\*  
13: gb\_est4.\*  
14: gb\_est5.\*  
15: em\_estfun.\*  
16: em\_estom.\*  
17: em\_gss\_hum.\*  
18: em\_gss\_inv.\*  
19: em\_gss\_pln.\*  
20: em\_gss\_vrt.\*  
21: em\_gss\_fun.\*  
22: em\_gss\_nam.\*  
23: em\_gss\_mus.\*  
24: em\_gss\_pro.\*  
25: em\_gss\_rod.\*  
26: em\_gss\_phg.\*  
27: em\_gss\_vrl.\*  
28: gb\_gasl.\*  
29: gb\_gse2.\*

ALIGNMENTS

RESULT 1  
BZ568926  
LOCUS  
DEFINITION  
pac82-164\_8156.x1 pac82-164 Pseudomonas aeruginosa genomic clone  
BZ568926\_8156, genomic survey sequence.  
ACCESSION  
BZ568926  
VERSION  
BZ568926.1  
KEYWORDS  
GI:27202739  
SOURCE  
Pseudomonas aeruginosa  
ORGANISM  
Pseudomonas aeruginosa  
Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;  
Pseudomonadaceae; Pseudomonas.  
REFERENCE  
1 (bases 1 to 1293)  
Spencer D.H., Raymond C.K., Smith, E.E., Sims, E.E., Hastings, M.,  
Burns, J.L., Kaul, R. and Olsen, M.V.  
AUTHORS  
Whole-Genome-Sequence variation among multiple isolates of  
Pseudomonas aeruginosa library  
TITLE  
J. Bacteriol. (2002) In press  
JOURNAL  
COMMENT  
Contact: Chris K. Raymond  
Genome Center  
University of Washington  
Box 352145, Seattle, WA 98105-2145, USA  
Tel: 2062216954  
Fax: 2066857244  
Email: craymond@u.washington.edu  
Class: shotgun.  
Location/Qualifiers

5 316.4 29.2 636 29 CC857250 ND1.101A2  
6 196 18.1 1026 28 BZ559351 pac82-164  
7 194.8 18.0 430 29 CC868241 NDL.44D16  
8 79 7.3 158 10 AM863470 MR3-SN000  
9 72 6.6 891 28 A2683582 ENTXK47PR  
10 69.8 6.4 332 28 A2578373 22b10 SHO  
11 66.8 6.2 816 28 A2535744 ENTQ25TR  
12 66.8 6.1 906 28 BH153606 ENTTS83TF  
13 65.8 6.2 529 14 CF143195 UI-HF-BR0  
14 64.6 6.0 890 28 BH146886 ENTXK48TF  
15 63 5.8 786 29 CG311108 OGYBG71TH  
16 63 5.8 818 14 CB833135 OSIEB11M  
17 63 5.8 818 14 CB833177 OSIEB11M  
18 63 5.8 834 14 CB649819 OSUNEBl3P  
19 62.6 5.8 406 14 CF142680 UI-HF-BR0  
20 61.8 5.7 480 9 AL828696 AL828696  
21 61.2 5.7 1300 11 AV106071 Zea mays  
22 61 5.6 597 28 BZ652916 OGAMC66TM  
23 61 5.6 654 14 CB863587 AZO1.107E  
24 61 5.6 721 13 BQ620395 TALr1164E  
25 60.8 5.6 682 12 BJ465499 BJ465499  
26 60.8 5.6 700 9 AL508300 AL508300  
27 60.6 5.6 789 29 CG214539 OGYBX83TH  
28 60.6 5.6 887 13 BQ731479 AGENCOURT  
29 60.2 5.6 526 9 AL825814 AL825814  
30 60.2 5.6 534 14 CB874750 AZO3.102P  
31 60.2 5.6 583 14 CA229130 SCAGFL302  
32 60.2 5.6 607 13 BQ620757 TALr1129F  
33 60.2 5.6 623 13 BQ620376 TALr1166D  
34 60.2 5.6 664 14 CB875475 AZO3.108E  
35 60.2 5.6 675 14 CB875233 AZO3.104J  
36 60.2 5.6 687 14 CB874084 AZO3.101F  
37 60.2 5.6 695 14 CB863246 AZO1.106C  
38 60.2 5.6 698 14 CB863803 AZO1.108A  
39 60.2 5.6 699 14 CB874586 AZO3.102I  
40 60.2 5.6 718 13 BQ789402 WHEA160.G  
41 59.4 5.5 587 14 CA220741 SCBGF405  
42 59.4 5.5 625 14 CD219591 C C C C C  
43 59.4 5.5 658 14 CD222026 CCC1.1.F0  
44 59.4 5.5 670 14 CD227954 CCC1.56.H  
45 59.4 5.5 683 14 CF847209 p8HB042SD

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	629.4	58.1	1293	28	BZ568926
2	580.8	53.6	1253	28	BZ579253
3	564	52.1	754	28	BZ550302
4	324.4	30.0	590	28	CC140622

ORGANISM	Pseudomonas aeruginosa	
	Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;	
	Pseudomonadaceae; Pseudomonas.	
	1 (bases 1 to 1253)	
REFERENCE	Spencer, D.H., Raymond, C.K., Smith, E.E., Sims, E.E., Hastings, M.,	
	Burns, J.L., Kaul, R. and Olsen, M.V.	
AUTHORS	Whole-genome-Sequence variation among multiple isolates of	
	Pseudomonas aeruginosa library	
TITLE	J. Bacteriol. (2002) In press	
	Contact: Chris K. Raymond	
JOURNAL	Genome Center	
	University of Washington	
COMMENT	Box 352145, Seattle, WA 98105-2145, USA	
	Tel: 2062216954	
	Fax: 2066857244	
	Email: craymond@u.washington.edu	
	Class: shotgun.	
	Location/Qualifiers	
FEATURES	1..1253	
	/organism="Pseudomonas aeruginosa"	
source	/mol_type="genomic DNA"	
	/strain="MSH"	
	/db_xref="taxon:287"	
	/clone="msh2_6268"	
	/note="Environmental isolate. Whole genomic shotgun	
	library."	
ORIGIN	Query Match 53.6%; Score 580.8; DB 28; Length 1253;	
	Best Local Similarity 95.9%; Pred. No. 3.4e-79;	
	Matches 636; Conservative 0; Mismatches 22; Indels 5; Gaps 4;	
	QY 260 CGTGTGTTGGCGGATCTTTCCAAACCGCTAGTGTGGGTAGTGTGTTACCTTCG 319	
Db	97 CCTCTGTGGGGGATCTTTCCAAACCGCTAGTGTGGGTAGTGTGTTACCTTCG 156	
	QY 320 TGTTCGTGTCATCGGTGGGTAGAGAGATCCGCAAGGTGATCGAGAAGACTCCGTC 379	
QY	157 TGTTCGTGTCATCGGTGGGTGGGAGATCTTCCAAACCGCTAGTGTGGGTAGTGTGTTACCTTCG 216	
	QY 380 GCTGTCGAGCGCTGGAAGTACTTCTGCGAGTCTGTTTCGGCATCGCGCGCGCGCTGT 439	
Db	217 GCTGTCGAGCGCTGGAAGTACTTCTGCGAGTCTGTTTCGGCATCGCGCGCGCGCTGT 276	
	QY 440 TCCTCTACATGATCGCGAAGACCCGATCGAGACACCCCTGATCGTGGCGATCGTGAAGA 499	
Db	277 TCCTCTACATGATCGCGAAGACCCGATCGAGACACCCCTGATCGTGGCGATCGTGAAGA 336	
	QY 500 GGTGCGAGATCCAGTTGGGCATCTTCTGTCGTCTCTGACCTTCTGTCATCGTCGGCT 559	
Db	337 GGTGCGAGATCCAGTTGGGCATCTTCTGTCGTCTCTGACCTTCTGTCATCGTCGGCT 396	
	QY 560 CGAGCAATCGGTGAACCTCACCGACGGTCTCGACGGCTGGCGATCATGCGACGGTAA 619	
Db	397 CGAGCAATCGGTGAACCTCACCGACGGTCTCGACGGCTGGCGATCATGCGACGGTAA 456	
	QY 620 TGGTTGCGCGCGCTGGGCATCTTCTGTCATCTCTCGGCAACCTGAAGTTTCGCGAGT 679	
Db	457 TGGTTGCGCGCGCTGGGCATCTTCTGTCATCTCTCGGCAACCTGAAGTTTCGCGAGT 516	
	QY 680 ACCTGCTGATTCCTCAACGTTACCGCGCGCGCGAGCTGATCGTGTCTCGCGCGCGCTGG 739	
Db	517 ACCTGCTGATTCCTCAACGTTACCGCGCGCGCGAGCTGATCGTGTCTCGCGCGCGCTGG 576	
	QY 740 TGGCGCGCGCGCTGGCTTCTCTGTTTCAACATCTATCCGCGCGAGTCTTCATGGGCG 799	
Db	577 TGGCGCGCGCGCTGGCTTCTCTGTTTCAACATCTATCCGCGCGAGTCTTCATGGGCG 636	
	QY 800 ACCTGCGCGCGCTGGCTGGCGCGCGCTGGCGACCATTCGC-GGTGATCGTGGCGCG 858	
Db	637 ACCTGCGCGCGCTGGCTGGCGCGCGCTGGCGACCATTCGC-GGTGATCGTGGCGCG 696	

source	1..1293	
	/organism="Pseudomonas aeruginosa"	
	/mol_type="genomic DNA"	
	/strain="2-164"	
	/db_xref="taxon:287"	
	/clone="pacs2-164_8156"	
	/clone_lib="pacs2-164"	
	/note="Clinical isolate 2-164 Whole genomic shotgun	
library."		
ORIGIN	Query Match 58.1%; Score 629.4; DB 28; Length 1293;	
	Best Local Similarity 97.2%; Pred. No. 1.2e-86;	
	Matches 650; Conservative 0; Mismatches 18; Indels 1; Gaps 1;	
	QY 405 CTGGCAGTCGGTTCGGCATCGCGCGCGCGTTCCTTACATGATCGCGCAACCC 464	
Db	90 CTGGCAGTCGGTTCGGCATCGCGCGCGCGTTCCTTACATGATCGCGCAACCC 149	
	QY 465 GATCGAGACCCCTGATCGTGGCATCTGCTGAAGAGCTGAGATCCAGTTGGGCATCTT 524	
Db	150 GATCGAGACCCCTGATCGTGGCATCTGCTGAAGAGCTGAGATCCAGTTGGGCATCTT 209	
	QY 525 CTTGCTGTCTGACCTACTTCTGTCATCGTTCGGTTCGAGCAATCGGTGAACCTTCACCGA 584	
Db	210 CTTGCTGTCTGACCTACTTCTGTCATCGTTCGGTTCGAGCAATCGGTGAACCTTCACCGA 269	
	QY 585 CGTCTCACGGCGCTGGCATCATGCCGAGCTAATGTTCCGCGCGCGCTGGGCATCTT 644	
Db	270 CGGTCTCACGGCGCTGGCATCATGCCGAGCTAATGTTCCGCGCGCGCTGGGCATCTT 329	
	QY 645 CTGTACTCTGTGGGCAACGTAAGTTCGCCGAGTACTGTCTGATTCCTCAACGTTACCGGG 704	
Db	330 CTGTACTCTGTGGGCAACGTAAGTTCGCCGAGTACTGTCTGATTCCTCAACGTTACCGGG 389	
	QY 705 CGCGCGGAGCTGATCGTTCGCGCGCGCTGTCGGCGCGCGCTGGGCATCTTCTGCTG 764	
Db	390 CGCGCGGAGCTGATCGTTCGCGCGCGCTGTCGGCGCGCGCTGGGCATCTTCTGCTG 449	
	QY 765 GTTCAACACCTATCGCGCGAGTCTTCATCGGCGAGCTGCGCGCGCTGGCGCTGGCGCG 824	
Db	450 GTTCAACACCTATCGCGCGAGTCTTCATCGGCGAGCTGCGCGCGCTGGCGCTGGCGCG 509	
	QY 825 CGCGTGGGACCATCGCGGATCGTTCGCGCGAGTCTGTCGGCGAGATCGTTCATCATGGT-G 883	
Db	510 CGCGTGGGACCATCGCGGATCGTTCGCGCGAGTCTGTCGGCGAGTCTGTCATCATGGTGG 569	
	QY 884 GGGTGTTCCTCATGAAACCTCTCGGTGATGATCCAGTCTGCTTCTTCAAGCTGACCG 943	
Db	570 GGGTGTTCCTCATGAAACCTCTCGGTGATGATCCAGTCTGCTTCTTCAAGCTGACCG 629	
	QY 944 GACGCGCGCTTTCGTTAGCGCGGATCCATACCATTTTCGAACTGAAGGCTGGCGCG 1003	
Db	630 GACGCGCGCTTTCGTTAGCGCGGATCCATACCATTTTCGAACTGAAGGCTGGCGCG 689	
	QY 1004 ACCGCGCGTGTATCGTGGCTTCTGATCATCACCGTATCCTGCTGATCGGCGCTCG 1063	
Db	690 AGCGCGCGTGTATCGTGGCTTCTTATCATCATCCGTTGATCTGTTGCTGATCGGCGCTC 749	
	QY 1064 CCACCTTGA 1072	
Db	750 GCCACTTGA 758	
RESULT 2	BZ579253	
	LOCUS	
DEFINITION	msh2_6268.x1 msh Pseudomonas aeruginosa genomic clone msh2_6268,	
	genomic survey sequence.	
ACCESSION	BZ579253	
	VERSION	
KEYWORDS	BZ579253.1 GI:27214314	
	GSS.	
SOURCE	Pseudomonas aeruginosa	

```

QY 859 G-AGATCGTGC-TGTTTCATCATGGGTGGG--TGTTGTCATGGAACCCCTCTCGGTGAT 914
Db 697 GAANATCGGTGTTGTTTCATTCATTTATGGTGGGTGTTGTCGTCATGGAACCCCTCTCGGT 756
QY 915 GAT 917
Db 757 GAT 759

RESULT 3
BZ550302/c
LOCUS
DEFINITION
pacal-60_2592.y2 pacal-60 Pseudomonas aeruginosa genomic clone
ACCESSION
BZ550302
VERSION
BZ550302.1 GI:27153883
KEYWORDS
GSS.
SOURCE
Pseudomonas aeruginosa
ORGANISM
Pseudomonas aeruginosa
Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
Pseudomonadaceae; Pseudomonas.
REFERENCE
1 (bases 1 to 754)
Spencer,D.H., Raymond,C.K., Smith,E.E., Sims,E.E., Hastings,M.,
Burns,J.L., Kaul,R. and Olsen,M.V.
Whole-Genome-Sequence variation among multiple isolates of
Pseudomonas aeruginosa library
J. Bacteriol. (2002) In press
Contact: Chris K. Raymond
Genome Center
University of Washington
Box 352145, Seattle, WA 98105-2145, USA
Tel: 2062216954
Fax: 2066857244
Email: craymond@u.washington.edu
Class: shotgun.
FEATURES
Source
Location/Qualifiers
1..754
/organism="Pseudomonas aeruginosa"
/mol_type="genomic DNA"
/strain="1-60"
/db_xref="taxon:287"
/clone="pacal-60_2592"
/clone_lib="pacal-60"
/note="clinical isolate 1-60 Whole genomic shotgun
library."
ORIGIN
Query Match 52.1%; Score 564; DB 28; Length 754;
Best Local Similarity 88.5%; Pred.No.1.1e-76;
Matches 634; Conservative 0; Mismatches 80; Indels 2; Gaps 2;

QY 205 AAGGGCACCCGACCATGGCGCGCCCTGATCTTACCGCATAGCCATCAGCAGCGTG 264
Db 753 AAAGGGCATCTGACCATGGCGCGCCCTAGATCC-TACCCCATAGCCATCAGCCGTTG 695
QY 265 CTGTGGCGGATCTTCCAAACGCTACGTTGGGTAGTCTGTTACCTGCTGTTTC 324
Db 694 CTGTGGCGGATCTTTCACACCGTACGTTGGTGAAGTCTGGTGGAAACCTAGCTGATC 635
QY 325 GTTGCCATCGGCTGGGTAGACCACTACCCCAAGGTGATCGAAGAAGTCTCCGTCGCTG 384
Db 634 GGAGCCACCGACTGGGTGGACAACCTACCTCAAGGTGATCGAAGAAGTCTCCATGCGCTG 575
QY 385 CCGAGCGGCTGGAGTACTCTTGGGAGTGGGTGTTGGGATCGGCGCGCGCGTGTCTTC 444
Db 574 CTTAGCTCTGGAAGTACTTCTTGGCAGTCGGTGTTCGGCATCGGCGCGCGCGTGTCTTC 515
QY 445 TACATGACTGCCGAACCCCGATCGAGACACCT-GATCGTGGCGATGCTGAAGAGGT 503
Db 514 TACATGACTGCCGAACCTCGATCGAGACCACTTAGATCTGTCGATGCTGAAGAGGT 455
QY 504 CGAGATCCAGTTGGGCATCTTCTTGTGTTCTGACTTACTTCTGTCATCTGTCGGTCCAG 563

```

```

Db 454 CGAGATCCAGTTGGGCATCTTCTTCTGGTCTGACTTACTTCTGTCATGTCGTTCCGAG 395
QY 564 CAATGGCGTGAACCTCACCGACGTTCTCGACGCTTGGCGATCATCGCAGTAATGTT 623
Db 394 CAATGCAGTGAACCTCACCGACGTTCTCGACGCTTGGCGATCATCGCAGTAATGTT 335
QY 624 TGGCGGCGCTGGGCATCTTCTTCTGCTACCTGTCGGGCAAGTGAAGTTCGCCGATACCT 683
Db 334 TGGCGGCGCTGGGCATCTTCTGCTACCTGTCGGGCAAGTGAAGTTCGCCGATACCT 275
QY 684 GCTGATTCCCAAGTACCGGCGCTGGGCGAGCTGATCTGTTCTTGGCGCGCTGTTGTCGG 743
Db 274 GTTGATTCCCTAGTACCGGCGCTGGGCGAGCTGATCTGTTCTGCGCGCATGTTGTCGG 215
QY 744 CGCGGCGCTGCGCTTCTTCTGTTTCAACCTATCGGCGCAGTCTTCTATGGCGACGT 803
Db 214 CGCGGCGCTGCGCTTCTTCTGTTTCAACCTATCGGCGCAGTCTTCTATGGCGACGT 155
QY 804 CGGCGGCTGCGCTGCGGCGCTGGGCGAGCTGATCTGTTCTGCGCGCTGATCTGCGCGCAGAGAT 863
Db 154 CGGCGGCTGCGCTGCGGCGCTGGGCGAGCTGATCTGTTCTGCGCGCTGATCTGCGCGCAGAGAGG 95
QY 864 GTGCTCTTTCATCATGGGTGGGTGTTGTCATGGAACCTCTTCTGCGTGAATGCC 919
Db 94 ATCCTCATAGTTCTAGAGCGCGCGCTCTCTCGATGAGTCTCTGTTTATGTTCC 39

RESULT 4
CC140622
LOCUS
DEFINITION
NDL.67G2.SP6 Notre Dame Liverpool Aedes aegypti genomic clone
ACCESSION
CC140622
VERSION
CC140622.1 GI:30009677
KEYWORDS
GSS.
SOURCE
Aedes aegypti (yellow fever mosquito)
ORGANISM
Aedes aegypti
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Aedes;
Stegomyia.
REFERENCE
1 (bases 1 to 590)
Loftus,B., Shetty,J., Knudson,D. and Severson,D.
BAC end sequencing of Aedes aegypti
Unpublished (2003)
Contact: Brendan Loftus
Department of Eukaryotic Genomics
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-3543
Fax: 301-838-0208
Email: enta@tigr.org
Library was provided by David Severson
Seq primer: SP6
Class: BAC ends.
FEATURES
Source
Location/Qualifiers
1..590
/organism="Aedes aegypti"
/mol_type="genomic DNA"
/strain="liverpool"
/db_xref="taxon:7159"
/clone="NDL.67G2"
/clone_lib="Notre Dame Liverpool"
/note="vector: PECBAC1, Site 1: Hind III; The library was
prepared from whole body tissue of newly hatched L1 larvae
by David Severson at the University of Notre Dame and
Hongbin Zhang"
ORIGIN
Query Match 30.0%; Score 324.4; DB 28; Length 590;
Best Local Similarity 73.1%; Pred.No.4.9e-40;
Matches 431; Conservative 0; Mismatches 156; Indels 3; Gaps 1;

QY 271 GCGGATCTTCCAAACCGCTACGTGTGGTAGTCTGTCGTTACCTGCTGTTGCGTGCC 330

```

[illegible]

RESULT 5	CC857250	636 bp	DNA	linear	GSS 24-JUL-2003
LOCUS	ND1.101A24.T7				
DEFINITION	Notre Dame Liverpool Aedes aegypti genomic clone				
	Notre Dame Liverpool-101A24.				
	genomic survey sequence.				

```

REFERENCE
AUTHORS      Lotius,B., Shetty,J., Knudson,D. and Severson,D.
TITLE        BAC end sequencing of Aedes aegypti
JOURNAL      Unpublished (2003)
COMMENT      Other GSSs: ND.L101A24.SP6
              Contact: Brendan Loftus
              Department of Eukaryotic Genomics
              TIGR
              9712 Medical Center Drive, Rockville, MD 20850, USA
              Tel: 301-338-3543
              Fax: 301-338-0208
              Email: enta@tigr.org
              Library was provided by David Severson
              Seq primer: T7
              Class: BAC ends.
              Location/Qualifiers
                  1..636
                     /organism="Aedes aegypti"
                     /mol_type="genomic DNA"
                     /strain="livernool"
FEATURES
source

```

FEATURES  
SOURCE

```
/db_xref="taxon:7159"  
/clone="NotreDame Liverpool-101824"  
/clone_lib="Notre Dame Liverpool"  
/note=vector: PECAC1, Site_1: Hind III; The library was  
prepared from whole body tissue of newly hatched L1 larvae  
by David Severson at the University of Notre Dame and  
Hongbin Zhang
```

ORIGIN	Query Match	29.2%; Score 316.4; DB 29; Length 636;
	Best Local Similarity	72.1%; Pred. No. 8.3e-39;
	Matches 454; Conservative 0;	Mismatches 171; Indels 5; Gaps 3
Qy	208	GGCACCCCGACATGGGCGCCCTGATCCTTACCGCATAGCCATCAGCACGCTGCTG 267
Db	7	GGCAGCGGACATGGGCGGTTCTGCTGATCTGCTCACCATCACCTGTCGGGTGATG 66
Qy	268	TGGCGGAGTCWTTCCAAACCGCTACGTGTGGGTAGTCTGGTGTACCTGCTGTTCGGT 327
Db	67	TGGCGGACCTTGGCAACCGCTACGTGTGGCTGCTGCGGTGATGTGTGCTTCGCG 126
Qy	328	GCATCGCTGGGTAGACGACTACCGCAAGGTGATCAGAGAAGATCCCGTGGCTGCG 387
Db	127	GCATCGGCTGGTATGACATCTGGATCAGATGTTGCGCGTGACCCGACGCGCTGAAG 186
Qy	388	AGCCGCTGGAAAGTATTTCTGSCAGTCTGGTGTTCGGCATCGCGCCGCGCTGTTCCTTAC 447
Db	187	TCGCGCTGGAAAGTACCTGCTCAGTCTGATCTTCGGCTGCGCTGCGGCGCTGTTCCTGTTC 246
Qy	448	ATGACTGCGGAAACCCCGATCGAACACACCTGATGTGCCGATGCTGAAGAGCTTCGAG 507
Db	247	TACAGGCGGAGTGGCGGACGCGCTGACCTTCTACATCCGATGTTCAAGTCGGTTCGG 306
Qy	508	ATCCAGTTG--GGCATCTTTCTTGGTGGTCTGACCTTCTGTCATCTCGGCTCGAGC 564
Db	307	CTCCGCTGGCGCGCATCGGCTTCGTGGCCATCGCCTACTTCTGGATCGTCGCTTCTCC 366
Qy	565	AATGGGTGAACCTACCGAGGCTCTGACGGCTGGCGATCATCGGACGGTATGTT 624
Db	367	AACGGGTGAACCTACCGAGGCTTTCAGCGGCTGGCGATCATGCCACCGTGTGTTG 426
Qy	625	GCGGCGGCTGGGATCTTCTGTACTGTTCGGGCAACGTGAAGTTCGCCGAGTACCTTG 684
Db	427	GCCTGGCGCTGGGGTGTTTCGCTACGCTTCGGGCAACGTGGTGTTCGCCAATACCTG 486
Qy	685	CTGATPCCAAAGTACCGGGCGCGGAGCT-GATCGTGTTCGCGCGCGCTGGTCGG 743
Db	487	CAGATCCCGAGATCCCGGGGCGCGGAGCTGTTTCATCATCTGTGGGCGCATCGCCG 546
Qy	744	CGC-CGGGCTCGGCTTCCTCTGTGTTCAACACTATCCGGCGCAGGTCTTCATGGGCGAG 802
Db	547	CGCGGGGCTGGGCTTCTCTGTGTTCAACACTACCGACCATGTGTTCTCATGGGCGACA 606
Qy	803	TCGGCGGCTGCGCTGGGCGCGCGGTGG 832
Db	607	TCGGCGCACTGCACTGGGTGCGGTGTTGG 636

RESULT 6	BZ559351	1026 bp	DNA	linear	GSS 17-DEC-2002
LOCUS	BZ559351				
DEFINITION	pac62-164_1435.s1 pac62-164 Pseudomonas aeruginosa genomic clone				
	pac62-164_1435, genomic survey sequence.				
ACCESSION	BZ559351				
VERSION	BZ559351.1	GI:27175535			
KEYWORDS	GSS.				
SOURCE	Pseudomonas aeruginosa				
ORGANISM	Pseudomonas aeruginosa				
	Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;				
	Pseudomonadaceae; Pseudomonas.				
REFERENCE	1 (bases 1 to 1026)				
AUTHORS	Spencer, D.H., Raymond, C.K., Smith, E.E., Hastings, M., Burns, J.L., Kaul, R. and Olsen, M.V.				

TITLE Whole-Genome-Sequence variation among multiple isolates of  
Pseudomonas aeruginosa library  
J. Bacteriol. (2002) In press  
COMMENT Contact: Chris K. Raymond  
Genome Center  
University of Washington  
Box 352145, Seattle, WA 98105-2145, USA  
Tel: 2062216954  
Fax: 2066857244  
Email: craymond@u.washington.edu  
Class: shotgun.

FEATURES

source  
1..1026  
/organism="Pseudomonas aeruginosa"  
/mol\_type="genomic DNA"  
/strain="2-164"  
/db\_xref="taxon:287"  
/clone="pacs2-164.1435"  
/clone\_lib="pacs2-164"  
/note="Clinical isolate 2-164 Whole genomic shotgun  
library."

ORIGIN

Query Match 18.1%; Score 196; DB 28; Length 1026;  
Best Local Similarity 97.2%; Pred. No. 2.5e-20;  
Matches 210; Conservative 0; Mismatches 5; Indels 1; Gaps 1;  
QY 1 ATGCTCTCTGCTGCGCGAATACCTGCAACAGTTCTACAGGGCTTCGGGCTTCACAG 60  
Db 440 ATGCTCTCTGCTGCGCGAATACCTGCAACAGTTCTACAGGGCTTCGGGCTTCACAG 499  
QY 61 TACCTGACCTCGGGGCAATCTCAGCGGCTCACCAGGCTGCTCGCTGCGTGGGCTG 120  
Db 500 TACCTGACCTCGGGGCAATCTCAGCGGCTCACCAGGCTGCTCGCTGCGTGGGCTG 559  
QY 121 GGGCCCTGATCGTACCTGTCAGATGCCAGATGCCAGCGCCGTCGCCAACGAC 180  
Db 560 GGGCCCTGATCGTACCTGTCAGATGCCAGATGCCAGCGCCGTCGCCAACGAC 619  
QY 181 GTTCGCACTCGCACT-CTCGAAGAGGGCACCC 215  
Db 620 GTTCGCACTCGCACTCTGTCGAAGAAGGGCACCC 655

RESULT 7

CC868241 430 bp DNA linear GSS 24-JUL-2003  
LOCUS NDL44D16.SP6 Notre Dame Liverpool Aedes aegypti genomic clone  
DEFINITION Notre Dame Liverpool-44D16, genomic survey sequence.  
ACCESSION CC868241  
VERSION CC868241.1 GI:33228251  
KEYWORDS GSS.  
SOURCE Aedes aegypti (yellow fever mosquito)  
ORGANISM Aedes aegypti  
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Aedes;  
Stegomyia.

REFERENCE

1 (bases 1 to 430)  
Loftus, B., Shetty, J., Knudson, D. and Severson, D.  
BAC end sequencing of Aedes aegypti  
Unpublished (2003)  
Other\_GSSs: NDL44D16.T7  
Contact: Brendan Loftus  
Department of Eukaryotic Genomics  
TIGR  
9712 Medical Center Drive, Rockville, MD 20850, USA  
Tel: 301-838-3543  
Fax: 301-838-0208  
Email: entaetigr.org  
Library was provided by David Severson  
Seq primer: SP6  
Class: BAC ends.  
Location/Qualifiers

FEATURES

source

1..430  
/organism="Aedes aegypti"  
/mol\_type="genomic DNA"  
/strain="liverpool"  
/db\_xref="taxon:7159"  
/clone="Notre Dame Liverpool-44D16"  
/clone\_lib="Notre Dame Liverpool"  
/note="Vector: pBAC1; Site 1: Hind III; The library was  
prepared from whole body tissue of newly hatched 1st larvae  
by David Severson at the University of Notre Dame and  
Hongbin Zhang"

ORIGIN

Query Match 18.0%; Score 194.8; DB 29; Length 430;  
Best Local Similarity 67.4%; Pred. No. 3.1e-20;  
Matches 290; Conservative 0; Mismatches 137; Indels 3; Gaps 1;  
QY 176 ACGAGGTCGCGAGTCGACCTGTCGAAGAAGGCGCACCCGACCATGGGGCGCCCTGA 235  
Db 1 AGAAGGTCGCGAGACCCATTCTCCAGGCCGCGACCCCTTCATGGGGGTTCTGCTGA 60  
QY 236 TCCTTACCGCCATAGCCATCAGCACGCTGCTGTGGGGGATCTTCCAAACCGTACTGTGT 295  
Db 61 TCCTGCTCACCATCACCTGCTGCTGCTGATGTGGGCGACCTGCGCAACCGCTACTGTGT 120  
QY 296 GGGTAGTCTGCTGCTTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 355  
Db 121 GGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180  
QY 356 AGGTGATCGAAGAAGAACTCCCGTGGCTGCGAGCGCTGGAAGTACTTCTGGCAGTCGG 415  
Db 181 AGATCGTCCGCGCTGACCCGAAAGCTGAGTGGCTGGAAGTACTTCTGCGAGTCGA 240  
QY 416 TGTTCGCACTCGGGCGCGCTGCTTCTCTACATGATGCGGAACCCCGATCGAGACCA 475  
Db 241 TCTTCGCGCTGGCTGGGGCTGCTTCTGCTTACAGCGCCGACGTCGCCGCGAGCGTGA 300  
QY 476 CCCTGATGCTGCGGATGCTGCAAGAGGCTCGAGATCCAGTTG---GGCATCTTCTTCGTGG 532  
Db 301 CCTTCTACATCCCGATGTTCAAGTCGCTGCGCTGCGCTGGCGGATCGGCTTCGTGG 360  
QY 533 TCCTGACCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 592  
Db 361 CCATCGCTACTTCTGATGCTGCTGCTTCCCAAGCGCTGAACCTGACCGCGGCTTG 420  
QY 593 ACGGCTGGC 602  
Db 421 ACGGCTGGC 430

RESULT 8

AW863470 158 bp mRNA linear EST 22-MAY-2000  
LOCUS MR3-SN0009-100400-206-e02 SN0009 Homo sapiens cDNA, mRNA sequence.  
DEFINITION AW863470  
ACCESSION AW863470  
VERSION AW863470.1 GI:7997520  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 158)  
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,  
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,  
Goldman, G.H., Carvalho, A.F., Matsuoka, A., Baia, G.S., Simpson, D.H.,  
Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V.,  
O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and  
Simpson, A.J.  
Shotgun sequencing of the human transcriptome with ORF expressed  
sequence tags  
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
MEDLINE 20202663  
PUBMED 10737800

COMMENT

Contact: Simpson A.J.G.  
Laboratory of Cancer Genetics  
Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01508-010, Sao Paulo-SP,  
Brazil  
Tel: +55-11-2704922  
Fax: +55-11-2707001  
Email: asimpson@ludwig.org.br  
This sequence was derived from the FAPESP/LICR Human Cancer Genome  
Project. This entry can be seen in the following URL  
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=st2=MR3-SN0009-100  
400-206-e02&k3=2000-04-10&t4=1)  
Seq primer: puc 18 forward  
High quality sequence start: 27  
High quality sequence stop: 158.  
Location/Qualifiers  
1..158  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/dev\_stage="Adult"  
/clone\_lib="SN0009"  
/note="Organ: stomach normal; Vector: puc18; Site 1: SmaI;  
Site 2: SmaI; A mini-library was made by cloning products  
derived from ORESSES PCR (O.S. Letters Patent application  
No. 196,716 - Ludwig Institute for Cancer Research)  
profiles into the puc 18 vector. Reverse transcription of  
tissue mRNA and cDNA amplification were performed under  
low stringency conditions."

FEATURES

source

ORIGIN

Query Match 7.3%; Score 79; DB 10; Length 158;  
Best Local Similarity 77.7%; Pred. No. 0.013;  
Matches 108; Conservative 0; Mismatches 30; Indels 1; Gaps 1;  
QY 778 CCGGCGCAGGTTTCATGGCGACGTCGGCGGCTGGCGGCGCGGCTGGCGAC 837  
Db 17 CTTGGCCAGGTTTCATGGCGATTCGGCGCCCTGGCCCTCGGGTGCCTGGCGAC 76  
QY 838 ATCGCGGTGATCGTGGCAGGA-GATCGTGTTCATCATGGTGGGGTTCGTCAT 896  
Db 77 ATCGCGGTATCACCAGGAGATCTGTCGCGCATGTCGTCGCGCGGCTTCGTTGC 136  
QY 897 GGAACCCCTTCGGTGATG 915  
Db 137 CGAAGCGTTCGGTAATG 155

RESULT 9

AZ683582/c  
LOCUS ENTXK47R Entamoeba histolytica Sheared DNA linear GSS 14-DEC-2000  
DEFINITION genomic, genomic survey sequence.  
ACCESSION AZ683582.1 GI:11820728  
VERSION AZ683582.1  
KEYWORDS GSS.  
SOURCE Entamoeba histolytica  
ORGANISM Entamoeba histolytica  
REFERENCE Eukaryota; Entamoebidae; Entamoeba.  
1 (bases 1 to 891)  
AUTHORS Lofthus.S., Van Aken.S. and Fraser.C.  
TITLE Determination of clone end sequences from Entamoeba histolytica  
HMI:IMSS sheared DNA library  
JOURNAL Unpublished (2000)  
COMMENT Contact: Brendan J Loftus  
Department of Eukaryotic Genomics  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel: 301 838 0208  
Fax: 301 838 3543  
Email: bjloftus@igr.org  
Clones are derived from the Entamoeba histolytica HMI:IMSS sheared  
DNA library

Seq primer: M13-Reverse

Class: shotgun  
High quality sequence start: 16  
High quality sequence stop: 694.  
Location/Qualifiers  
1..891  
/organism="Entamoeba histolytica"  
/mol\_type="genomic DNA"  
/strain="HMI:IMSS"  
/db\_xref="taxon:5759"  
/clone\_lib="Entamoeba histolytica Sheared DNA"  
/note="Vector: pHOS1; Site 1: Bst I; Constructed at The  
Institute for Genomic Research (TIGR), Rockville, MD.  
Genomic DNA isolated from broth cultures of E. histolytica  
using a method described by Clark and Diamond (Clark,  
C.G., and Diamond, L.S. (1993) Entamoeba histolytica: a  
method for isolate identification. Exp. Parasitol.  
77:450.). The DNA was mechanically sheared to give a  
tight size distribution (~2 kb). The v + i method used for  
the library construction is described in detail in Smith,  
H.O. and Venter, J.C. (Making small insert libraries for  
whole genome shotgun sequencing projects. In Genome  
Sequencing: A Practical Approach, eds. M. Vaudin and B.  
Barell, Oxford University Press, 1999)."

FEATURES

source

ORIGIN

Query Match 6.6%; Score 72; DB 28; Length 891;  
Best Local Similarity 44.4%; Pred. No. 0.23;  
Matches 291; Conservative 0; Mismatches 365; Indels 0; Gaps 0;  
QY 415 GTGTCGGCATCGGCGCGCTGTTCTCTACATGACTGCGAAACCGCATCGAGAC 474  
Db 691 GTGCTTTCATGCTGCTTTCATGCTGCTTTCATGCTTCTTTCATGCTTCTTCATC 632  
QY 475 ACCTGATCGTCCGATGCTGAAGAGCGTCGAGATCCAGTGGGATCTTCTTCTGTTGTC 534  
Db 631 ATCATCATCTTCATCATCTTCTTCTTCATCATCGTCTTCTTCATCATCGTCTTCTTC 572  
QY 535 CTGACCTACTGTCATCGTCGCGCTCGAGCAATCGGTGAACCTCACGACGCTCTCGAC 594  
Db 571 ATCATGCTTCTTTCATCATCGTCTTCTTCATCATCGTCTTCTTCATCATCGTCTTCTTC 512  
QY 595 GGCTGGCGCATATGCGGAGGTAATGTTGCGGCGCGCTGGGATCTTCTGCTACCTG 654  
Db 511 ATCATGCTTCTTTCATCATCGTCTTCTTCATCATCGTCTTCTTCATCATCGTCTTCTTC 452  
QY 655 TCGGGCAAGTGAAGTTCGCGAGTACCTGCTGATTCCTCCAGCTACCGGCGCGCGGAG 714  
Db 451 ATCATGCTTCTTTCATCATCGTCTTCTTCATCATCGTCTTCTTCATCATCGTCTTCTTC 392  
QY 715 CTGATCGTGTTCGCGCGCTGGTGGCGCGCGCTCGGCTTCTCTGTTCAACACC 774  
Db 391 ATCATGCTTCTTAATCAATTCATCATCGTCTTCTTCATCATCGTCTTCTTCATCGTC 332  
QY 775 TATCGCGGAGGTCTTTCATGGGCGACGTCGGGCGGCTGGCGCTGGGCGCGCGGCGGCG 834  
Db 331 TTCTTCATCGTCTTCTTTCATCGTCTTCTTCATCATCGTCTTCTTCATCATCGTCTTCTTC 272  
QY 835 ACCATCGCGTATCGTGGCCAGGAGATCGTGTGTTTCATCATGGTGGGCTGTTCTGTC 894  
Db 271 TTCTTCATGCTTCTTTCATCATCATCATCTTCTTCATCATCATCATCATCTTCTTCATC 212  
QY 895 ATGGAACCTCTCGGTGATGATCCAGGTGCTTCTTCAAGTGAACCGGACCGCGCGCTC 954  
Db 211 ATCATCGTCTTTCATCATCATCTTCTTTCATCATCATCTTCTTCATCATCATCTTCTTC 152  
QY 955 TTCGGTATGGCGGATCCATCACCATTTCGAACGTTGAGGCGGCGGCGCGCGCTG 1014  
Db 151 ATCATCATCTTCTTTCATCATCATCTTCTTCATCATCATCTTCTTCATCATCATCTTCTTC 92  
QY 1015 ATCGTGGCTCTTGATCATCACCGGTGATCTCTGGTGTCTGATCGGCTCGCCACCTT 1070  
Db 91 ATCATGCTTCTTAAATTCATCATCTTTCATCATCATCATCATCATCTTTCATCATCAT 36

JOURNAL  
COMMENT

Unpublished (2000)  
Contact: Brendan J Loftus  
Department of Eukaryotic Genomics  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel: 301 838 0208  
Fax: 301 838 3543  
Email: b.loftus@tigr.org  
Clones are derived from the Entamoeba histolytica HM1:IMSS sheared  
DNA library  
Seq primer: M13-Reverse  
Class: Shotgun  
High quality sequence start: 36  
High quality sequence stop: 816.

## FEATURES

source

1..816

/organism="Entamoeba histolytica"  
/mol\_type="genomic DNA"  
/strain="HM1:IMSS"  
/db\_xref="taxon:5759"  
/clone\_lib="Entamoeba histolytica Sheared DNA"  
/notes="Vector: pHS1; Site 1: Bst I; Constructed at The  
Institute for Genomic Research (TIGR), Rockville, MD  
Genomic DNA isolated from broth cultures of E. histolytica  
using a method described by Clark and Diamond (Clark,  
C.G., and Diamond, L.S. (1993) Entamoeba histolytica: a  
method for isolate identification. Exp. Parasitol.  
77:450.). The DNA was mechanically sheared to give a  
tight size distribution (~2 kb). The v + i method used for  
the library construction is described in detail in Smith,  
H.O. and Venter, J.C. (Making small insert libraries for  
whole genome shotgun sequencing projects. In Genome  
Sequencing: A Practical Approach, eds. M. Vaudin and B.  
Barell, Oxford University Press, 1999)."

## ORIGIN

Query Match 6.2%; Score 66.8; DB 28; Length 816;

Best Local Similarity 43.0%; Pred. No. 1.4;

Matches 329; Conservative 0; Mismatches 437; Indels 0; Gaps 0;

QY	291	CGTGTGGGTAGTGTGCTGCTTACCCCTGCTGTTCGGTGCCATCGGTGGGTAGACGACTA	350
DB	48	CGTCTTCTTCATCGTCTTCTTCATCGTCTTTCATCGTCTTTCATCGTCTTTCAT	107
QY	351	CCGCAAGGTATCGAAGAACTCCCGTGCCTGCCGAGCGGTGGAGTACTTCTGGCA	410
DB	108	CGTGTCTTTCATCGTCTTTCATCGTCTTTCATCGTCTTTCATCGTCTTTCAT	167
QY	411	GTCGGTGTTCGGCATCGCGCGCGGCTTCTCTACATGACTGCCGAAACCCGATCGA	470
DB	168	CATCATCATCTTCATCATCTTCTTTCATCATCGTCTTTCATCATCGTCTTCTT	227
QY	471	GACCACCTGATCGTGCAGATGCTGAAGAGCGTGCAGATCCAGTTGGGCACTTCTTCGT	530
DB	228	CATCATCGTCTTCTTCATCATCGTCTTCTTCATCATCGTCTTCTTCATCATCGTCTTCTT	287
QY	531	GGTCTGACCTACTTGTGCTGCTGCGGCTCGAGCAATGCGGTGAACCTCACCGACGTCT	590
DB	288	CATCATCGTCTTCTTCATCATCGTCTTCTTCATCATCGTCTTCTTCATCATCGTCTTCTT	347
QY	591	CGACGGCTTGCGATCATGCCGAGCGGTAAATGGTTGCGCGCGCGTGGGCACTTCTTCGCTA	650
DB	348	CATCATCGTCTTCTTCATCATCGTCTTCTTCATCATCGTCTTCTTCATCATCGTCTTCTT	407
QY	651	CCTGTGGGCAACGTGAAGTTCCGCGAGTACCTGCTGATTCGCAACGTCGCGGCGCGG	710
DB	408	CATCATCGTCTTCTTCATCATCGTCTTCTTCATCATCGTCTTCTTCATCATCGTCTTCTT	467
QY	711	CGAGCTGATCGTGTTCGCGCGCGGTGGTCCGCGCGCGCTCGGCTTCCTCTGGTTCAA	770
DB	468	CGTCTTCTTCATCGTCTTCTTCATCGTCTTCTTCATCATCGTCTTCTTCATCATCGTCTTCTT	527
QY	771	CACCTATCCGGCGCAGGTCTTTCATGGCGGAGTTCGGCGCGCTGGCGCTGGCGCGCGCT	830

## RESULT 10

AZ578373/c

LOCUS

DEFINITION 22b10 Shot-gun genomic library of Rhizobium strain ANU265 Rhizobium  
sp. NGR234 genomic clone 22b10, genomic survey sequence.

ACCESSION AZ578373

VERSION AZ578373.1

KEYWORDS GSS.

SOURCE Rhizobium sp. NGR234

ORGANISM Rhizobium sp. NGR234

REFERENCE 1 (bases 1 to 332)  
Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
Rhizobiaceae; Rhizobium/Agrobacterium group; Rhizobium.

Viprey, V., Rosenthal, A., Broughton, W.J. and Perret, X.

Genetic snapshots of the Rhizobium species NGR234 genome

Genome Biol. 1 (6), RESEARCH0014 (2000)

21114532

11178268

Contact: Virginie Viprey

Laboratoire de Biologie Moleculaire des Plantes Superieures

University of Geneva

1 Chemin de l'Imperatrice, Chambesy/Geneva 1292, Switzerland

Tel: +44(0)1603450000

Fax: +44(0)1603450045

Email: virginie.viprey@bbsrc.ac.uk

Class: shotgun.

Location/Qualifiers

1..332

/organism="Rhizobium sp. NGR234"

/mol\_type="genomic DNA"

/strain="ANU265"

/db\_xref="taxon:394"

/clone\_lib="Shot-gun genomic library of Rhizobium strain

ANU265"

/notes="Vector: M13; derivative strain of NGR234 cured of

pNGR234a"

Query Match 6.4%; Score 69.8; DB 28; Length 332;

Best Local Similarity 71.3%; Pred. No. 0.39; Indels 0; Gaps 0;

Matches 92; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

## ORIGIN

QY	955	TTCCGTATGGCGCGATCCATACCATTCGAACTGAAAGGTGGCCGACCCGCGGTG	1014
DB	332	TTCCGTATGGCGCGCAATCCACCATCACTTCGAGAGAGGGCTGGACCGAGCAGGTG	273
QY	1015	ATCGTGGCTTCGGATCATCCGNGATCCTGCTGATCGGCTCCACCTTGAAG	1074
DB	272	GTGATCGCTTCGGATCATCCGCGGTCATCTGGCGATGTCGGTCTTCGACCCCTCAG	213
QY	1075	CTGCGTTGA	1083
DB	212	CTGCGGTAA	204

RESULT 11  
AZ535744  
LOCUS  
DEFINITION ENTQ25TR Entamoeba histolytica Sheared DNA Entamoeba histolytica  
genomic, genomic survey sequence.  
ACCESSION AZ535744  
VERSION AZ535744.1  
KEYWORDS GSS.  
SOURCE Entamoeba histolytica  
ORGANISM Entamoeba histolytica  
REFERENCE 1 (bases 1 to 816)  
Eukaryota; Entamoebidae; Entamoeba.  
Loftus B., Van Aken, S. and Fraser, C.  
Determination of clone end sequences from Entamoeba histolytica  
HM1:IMSS sheared DNA library

Db	528	CGTCTTCTTCATCGTCTCTTCATCGTCTTCTTCATCGTCTTCTTCATCGTCTTCATCATCAT	587
Qy	831	GGGACCAATCGCGGTGATCGTGGCCAGGAGATCGTCTTTCATCATGGTGGGTGTT	890
Db	588	CTTCTTTCATCATCATCGTCTTCATCATCATCGTCTTCATCATCATCTTCTTCAT	647
Qy	891	CGTCATGGAAACCTCTCGGTGATGATCCAGGTGCTTCTTCAAGTGCACCGAGCGCG	950
Db	648	CATCATCTTCTTCATCATCATCTTCTTCATCATCATCTTCTTCATCATCATGTCATCTT	707
Qy	951	CGTCTTCGGTATGGCGCGATGCATACCATTTGCACTGAAGGTGGCGGACCGCG	1010
Db	708	CATCATCATCTTCTTCATCATCATCTTTCATCATCGTCTTCTTAATTCAAATTCATCATCTG	767
Qy	1011	CGTGATCGTGGCTCTTGATCATCATCCGTGATCTCGTGGTGTATC	1056
Db	768	CATCATCATCTTTCATCATCATCTTCTTCATCATCGTCTTCATC	813

RESULT 12					
BH153606					
LOCUS	BH153606	906 bp	DNA	linear	GSS 24-SEP-2001
DEFINITION	ENTTS83TF Entamoeba histolytica Sheared DNA				
	genomic, genomic survey sequence.				

BH153606  
 BH153606.1 GI:15725323  
 GSS.  
 SOURCE  
 ORGANISM  
 Entamoeba histolytica  
 Entamoeba histolytica  
 Eukaryota; Entamoebidae; Entamoeba.  
 1 (bases 1 to 906)  
 Loftus, S., Wang, Z., Van Aken, S. and Fraser, C.  
 Determination of clone end sequences from Entamoeba histolytica  
 HM1:IMSS sheared DNA library (2001)  
 Unpublished (2001)  
 JOURNAL  
 COMMENT  
 Contact: Brendan J Loftus  
 Department of Eukaryotic Genomics  
 The Institute for Genomic Research  
 9712 Medical Center Dr., Rockville, MD 20850, USA  
 Tel: 301 838 0208  
 Fax: 301 838 3543  
 Email: b.loftus@igr.org  
 Clones are derived from the Entamoeba histolytica HM1:IMSS sheared  
 DNA library

```

FEATURES
source
high quality sequence stop: 733.
Location/Qualifiers
1..906
/organism="Entamoeba histolytica"
/mol_type="genomic DNA"
/strain="HM1:IMSS"
/db_xref="taxon:5759"
/clone_lib="Entamoeba histolytica Sheared DNA"
/notice=vector; pHOS1; Site 1; Bat I; Constructed at The
Institute for Genomic Research (TIGR), Rockville, MD.
Genomic DNA isolated from broth cultures of E. histolytica
using a method described by Clark and Diamond (Clark,
C.G., and Diamond, L.S. [1993] Entamoeba histolytica: a
method for isolate identification. Exp. Parasitol.
77:450.). The DNA was mechanically sheared to give a
tight size distribution (~2 kb). The v + i method used for
the library construction is described in detail in Smith,
H.O. and Venter, J.C. (Making small insert libraries for
whole genome shotgun sequencing projects. In Genome
Sequencing: A Practical Approach, eds. M. Vaudin and B.
Barrell. Oxford University Press, 1999)."

```

ORIGIN

Query Match 6.2%; Score 66.8; DB 38; Length 906;  
Best Local Similarity 43.3%; Pred. NO. 1.4;  
Barell, Oxford University Press, 1999).

Matches	314;	Conservative	0;	Mismatches	412;	Indels	0;	Gaps	0
Qy	232	CTGATCCTTTACGGCATAGCCATCAGACAGCTGTGTGGCGGATCTTTCCAAACGCGTAC	291						
Db	23	CTAATTCATTCATCATCGTCTTCTTCATCGTCTTCTTCATCGTCTTCTTCATCGTCTTC	82						
Qy	292	GTGTGGGTAGTGTGGTGTACCTCGTGTTCGGTGGCATCGGCTGGGTAGACGACATAC	351						
Db	83	TTTCATCGTCTTCTTCATCGTCTTCTTCATCGTCTTCTTCATCGTCTTCTTCATCGTCTTC	142						
Qy	352	CGCAAGGTGATCGAGAAGAACTCCCGTGGCTGCGAGCGGTGGAAGTACTTCTGGCAG	411						
Db	143	TTTCATCGTCTTCTTCATCGTCTTCTTCATCGTCTTCTTCATCGTCTTCTTCATCGTCTTC	202						
Qy	412	TGGGTGTTGGGATCGGCGCGCGTGTTCCTTCATAGATGCCGAAACCCGACPCGAG	471						
Db	203	TTTCATCGTCTTCTTCATCGTCTTCTTCATCGTCTTCTTCATCGTCTTCTTCATCGTCTTC	262						
Qy	472	ACGACGCTGATGTCGCGATGCTGAAGAGCGTCGAGATCCAGTTGGGCACTTCTTCGTG	531						
Db	263	TTTCATCATCTTCTTCATCGTCTTCTTCATCATCATCATCTTCATCATCTTCATCTTC	322						
Qy	532	GTCTCTGACCTACTTTCGTTCATCGTGGCTCGAGCAATGCGTGAACTCACCACGCGTCTC	591						
Db	323	ATCATCGTCTTCTTCATCATCATCATCATCATCATCATCATCATCATCATCATCATCTTC	382						
Qy	592	GACGCGCTGGGATCAGCCGACGATTAATGTTTCCGGGCGCGTGGGCATCTTCGTGCTAC	651						
Db	383	ATCATCGTCTTCTTCATCATCGTCTTCTTCATCATCGTCTTCTTCATCATCGTCTTCTTC	442						
Qy	652	CTGTGGGCAAGTGAAGTTGCGCGAGTACCTGCTGATTCGCCAAGTACGCGGCGCGCG	711						
Db	443	ATCATCGTCTTCTTCATCATCATCATCATCATCATCATCATCATCATCATCATCATCTTC	502						
Qy	712	GAGCTGATCGTGTCTGCGCGCGCTGGTGGGCGCGCGCTCGGCTTCCTCTGGTTTCAAC	771						
Db	503	ATCATCGTCTTCTTCATCATCATCATCATCATCATCATCATCATCATCATCATCATCTTA	562						
Qy	772	ACCTATCGCGCGAGGTTTTCATGGGCGAATGTCGGGCGCGTGGCGTGGGCGCGCGCGT	831						
Db	563	TTCAAATTTCATCATCGTCTTCTTCATCATCGTCTTTCATCATCGTCTTCTTCATCGTCTTC	622						
Qy	832	GGCACCATCGGGTATCGTGGCGCAGAGATCGTGTGTTTCATCATGGTGGGGTGTTC	891						
Db	623	TTTCATCGTCTTCTTCATCATCATCATCATCATCATCATCATCATCATCATCATCATCTTC	682						
Qy	892	GTTCATGAAACCTCTCGGTGATGATCCAGGTGCGTTCCTTTCATAGTGCACCGACGCGCG	951						
Db	683	CTCATCGGATCATCATCTTCTTCATCATCATCATCATCATCATCATCATCATCATCGCATCTTC	742						
Qy	952	GTCTTC	957						
Db	743	ATCATC	748						

RESULT 13	CF143195	529 bp	linear	EST 09-SEP-2003
LOCUS	CF143195			
DEFINITION	UI-HF-BR0p-aqz-b-02-0-UI.r1 NIH_MGC_52 Homo sapiens cDNA clone IMAGE:3102243 5', mRNA sequence.			
VERSION	CF143195			
ACCESSION	CF143195.1	GI:33258639		
KEYWORDS	EST.			
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
REFERENCE	1 (bases 1 to 529)			
AUTHORS	Bonaldo, M.F., Lennon, G. and Soares, M.B.			
TITLE	Normalization and subtraction: two approaches to facilitate gene discovery			
JOURNAL	Genome Res. 6 (9), 791-806 (1996)			
MEDLINE	97044477			

## PUBMED

8889548  
Contact: Soares, MB  
Coordinated Laboratory for Computational Genomics  
University of Iowa  
375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA  
Tel: 319 335 8250  
Fax: 319 335 9565  
Email: bento-soares@uiowa.edu

Tissue Procurement: Louis Staudt  
cDNA library preparation: Dr. M. Bento Soares, University of Iowa  
cDNA library Arrayed by: Dr. M. Bento Soares, University of Iowa  
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
Clone Distribution: Distribution information can be found at  
<http://genome.uiowa.edu/distribution/humanfl.html>  
Seq primer: ptx-5,  
Location/Qualifiers

FEATURES  
source

1. .529  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:3102243"  
/tissue\_type="lymph"  
/cell\_type="germinal center B cells"  
/cell\_line="MGC85"  
/lab\_host="DH10B (LTI)"  
/clone\_lib="NIH MGC 52"  
/note="Vector: p77T3-Pac; Site 1: NotI; Site 2: Eco RI;  
Constructed from size fractionated cytoplasmic mRNA  
(7.4-9.5Kb). Directionally cloned. Cells provided by  
Louis M. Staudt, Ph.D. Library preparation by Maria de  
Fatima Bonaldo, Ph.D. and M. Bento Soares, Ph.D."

## ORIGIN

Query Match 6.1%; Score 65.8; DB 14; Length 529;  
Best Local Similarity 59.3%; Pred. No. 1.8; Indels 0; Gaps 0;  
Matches 112; Conservative 0; Mismatches 77; Indels 0; Gaps 0;  
QY 670 TTGCGGAGTACTGCTGATGATCCACGACGACGCGCGCGGCGGAGTGATGCTGTCGC 729  
Db 47 TTCAGTGCCTACGCGCTGCTGGCTTGTGCGGCGACGCGGAAATGGCTGCTGCTGACG 106  
QY 730 GCGCGCTGTCGCGCGCGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 789  
Db 107 GTGCTGCTGCGCGCTGCGGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 166  
QY 790 TTGATGGGCGACGCTGCGCGCGCTGCGCGCTGCGCGCGCTGCGCGCGCTGCGCGCGCTG 849  
Db 167 TTGATGGGCGACGCTGCGCGCGCTGCGCGCGCTGCGCGCGCTGCGCGCGCTGCGCGCG 226  
QY 850 GTGCGGCGAG 858  
Db 227 GGTATCCGG 235

## RESULT 14

BH146886/c  
LOCUS BH146886 890 bp DNA linear GSS 27-AUG-2001  
DEFINITION ENTPEK487F Entamoeba histolytica Sheared DNA Entamoeba histolytica  
genomic, genomic survey sequence.

ACCESSION BH146886.1 GI:15302963

## KEYWORDS

SOURCE Entamoeba histolytica

## ORGANISM

Entamoeba histolytica

Eukaryota; Entamoebidae; Entamoeba.

Loftus, B., Wang, Z., Van Aken, S. and Fraser, C.

Determination of clone end sequences from Entamoeba histolytica

(bases 1 to 890)

HM1:IMSS sheared DNA library (2001)

Unpublished (2001)

Contact: Brendan J Loftus

Department of Eukaryotic Genomics

The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850, USA

Tel: 301 838 0208

Fax: 301 838 3543

Email: b1loftus@tigr.org

Clones are derived from the Entamoeba histolytica HM1:IMSS sheared

DNA library

Seq primer: M13-Forward

Class: shotgun

High quality sequence start: 17

High quality sequence stop: 880.

Location/Qualifiers

FEATURES  
source

1. .890

/organism="Entamoeba histolytica"

/mol\_type="genomic DNA"

/strain="HM1:IMSS"

/db\_xref="taxon:5759"

/clone\_lib="Entamoeba histolytica Sheared DNA"

/note="Vector: pHO81; Site 1: Bst I; Constructed at The

Institute for Genomic Research (TIGR), Rockville, MD.

Genomic DNA isolated from broth cultures of E. histolytica

using a method described by Clark and Diamond (Clark,

C.G., and Diamond, L.S. (1993) Entamoeba histolytica: a

method for isolate identification. Exp. Parasitol.

77:450.). The DNA was mechanically sheared to give a

tight size distribution (~2 kb). The v + i method used for

the library construction is described in detail in Smith,

H.O. and Venter, J.C. (Making small insert libraries for

whole genome shotgun sequencing projects. In Genome

Sequencing: A Practical Approach, eds. M. Vaudin and B.

Barell, Oxford University Press, 1999)."

## ORIGIN

Query Match 6.0%; Score 64.6; DB 28; Length 890;  
Best Local Similarity 43.1%; Pred. No. 3.1;  
Matches 313; Conservative 0; Mismatches 414; Indels 0; Gaps 0;  
QY 231 CCTGATCGTTACCGCATAGCATACAGCAGCTGCTGGCGGAGTTCACCGCTA 290  
Db 791 CATCATCGTCTTTCATCGTCTTTCATCGTCTTTCATCGTCTTTCATCGTCTT 732  
QY 291 CGTGTGGTGTGCTGCTGCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 350  
Db 731 CTTTCATCGTCTTTCATCGTCTTTCATCGTCTTTCATCGTCTTTCATCGTCTT 672  
QY 351 CCGCAAGGTGATCGAAGAACTCCCGTGGCTGCGAGCGGCTGGAGTACTTTCGCA 410  
Db 671 CTTTCATCGTCTTTCATCGTCTTTCATCGTCTTTCATCGTCTTTCATCGTCTT 612  
QY 411 GTCGGTGTGGCATCGCGCGCGGCTGCTTCTACATGACTGCCGAACCCGATCGA 470  
Db 611 CTTTCATCGTCTTTCATCGTCTTTCATCGTCTTTCATCGTCTTTCATCGTCTT 552  
QY 471 GACCACCTGATCGTGGCGATGCTGAAGAGCGTTCAGATCCAGTTGGGCTATCTTTCGT 530  
Db 551 CTTTCATCATCTTTCATCGTCTTTCATCATCATCTTTCATCATCTTTCATCTTCTT 492  
QY 531 GGTCTGACCTACTTCGTCATCGTCGCTCGAGCAATCGGTGAACCTCACCGAGTCT 590  
Db 491 CATCATCGTCTTTCATCATCGTCTTTCATCATCATCTTTCATCATCGTCTTCTT 432  
QY 591 CGAGGCGCTGGCGATCATGCGCGAGTAATGTTGCGCGCGCTGGGCTATCTTTCGTA 650  
Db 431 CATCATCGTCTTTCATCATCGTCTTTCATCATCATCTTTCATCATCGTCTTCTT 372  
QY 651 CCTGTGGGCAACGTGAAGTTGCGCGAGTACCTGCTGATTCGCCAAGTACCGGCGGG 710  
Db 371 CATCATCGTCTTTCATCATCGTCTTTCATCATCATCTTTCATCATCGTCTTCTT 312  
QY 711 CGAGTGTATGCTGTTCTGCGCGCGCTGGTGGCGCGCGCTCGGCTTCCTCTCGTTCAA 770  
Db 311 CATCATCGTCTTTCATCATCGTCTTTCATCATCATCTTTCATCATCGTCTTCTA 252  
QY 771 CACCTATCCGCGGAGGTCTTTCATGGGCGAGCTGGGCGCGCTGGCGCTGGCGCGCT 830

```

Db      251 ATTCAAATTCATCATCGTCTTCTTCATCATCGTCTTCTTCATCGTCTTCTTCATCGTCTT 192
Qy      831 GGGACCATCGCGGTGATCGTGGCCAGGAGATCGTCTTTCATCATCGGTGGGTGTT 890
Db      191 CTTTCATCGTCTTCTTCATCGTCTTCTTCATCGTCTTCTTCATCGTCTTCTTCATCGTCTT 132
Qy      891 CGTCATGAAACCCCTCTCGGTGATGATCCAGGTGCTTCCTTCAAGCTGACCGGAGCGCG 950
Db      131 CTTTCATCGTTCATCATCTTCTTCATCATCGTCTTCTTCATCATCGTCTTCTTCATCATCTT 72
Qy      951 CGTCTTC 957
Db      71 CATCATC 65

RESULT 15
LOCUS   CG3111108 786 bp DNA linear GSS 26-AUG-2003
DEFINITION OGBEGLITH ZM_0.7_1.5_KB Zea mays genomic clone ZMMBma0640K21,
            genomic survey sequence.
ACCESSION CG3111108
VERSION   CG3111108.1 GI:34228268
KEYWORDS GSS.
SOURCE   Zea mays
ORGANISM Zea mays
          Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
          Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
          clade; Panicoideae; Andropogoneae; Zea.
          1 (bases 1 to 786)
          Whitelaw C.A., Quackenbush, J., Van Aken, S., Utterback, T.,
          Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,
          Citek, R.W., Nurnberg, A., Robbins, D. and Lakey, N.
          Consortium for Maize Genomics
          Unpublished (2002)
          Other GSSs: OGYBG71rv
          Contact: Cathy Whitelaw
          TIGR
          9712 Medical Center Drive, Rockville, MD 20850, USA
          Tel: 301-838-5843
          Fax: 301-838-0208
          Email: whitelaw@tigr.org
          Seq primer: TR
          Class: sheared ends.
FEATURES             Location/Qualifiers
     source           1..786
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /strain="B73"
                     /db_xref="taxon:4577"
                     /clone="ZMMBma0640K21"
                     /clone_lib="ZM 0.7 1.5 KB"
                     /note="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb
                     methylation filtered genomic DNA library"

ORIGIN
Query Match      5.8%; Score 63; DB 29; Length 786;
Best Local Similarity 46.4%; Pred. No. 5.3;
Matches 242; Conservative 0; Mismatches 277; Indels 3; Gaps 1;

Qy      376 CGTGGCCCTGCCGAGCGCGTGAAGTACTTCTGGCAGTGGTTCGGCATCGCGCGCGCC 435
Db      759 COTCGGCTGCTCGTGGACGGGGTACCGCCCTCAGCTGGTGGTGGCGGTGGCGATCCCT 700
Qy      436 GTGTTCTCTACATGATCGCCGAACCCGATCGAGACCACTCTGATCGTCCGATGCTG 495
Db      699 TTTCTTCAGCAGGTACCGGGATCAACCGATCGGTTCTACGGCCCGGTGCTGCTGCGC 640
Qy      496 AAGACGTCGAGATCCAGTGGCATCTTCTCGTGGTCTGACCTACTTCGTCTCATGTC 555
Db      639 ACCATCGGCATGGCGAGAGCGCTGCTGCTGTGTCGGGTGGTGACGGCGCTGTTGGC 580
Qy      556 GGCTCGAGCAATCGGTGAACCTCACCGACCGGTCTTCGACGGCCTCGGATCATGCCGACG 615

```

Search completed: May 9, 2004, 15:45:41  
Job time : 3224 secs

```

Db      579 GTGGCCTCCACCTCCGCGTCCATGCTGCGCG- - - TGGACCGCTTTGGGCGCGCACGCTG 523
Qy      616 GTAATGGTTGCCCGCGCTGGGCATCTTCTGCTACCTGTGCGGCAAGCTGAAGTTGCCC 675
Db      522 TTCCTGGCCGCGCGCGCGCAGATGCTGGCGTGCAGGTGCTGATCGGCGCATCATGCGC 463
Qy      676 GAGTACCTGCTGATTCCCAACGTACCGGGCGCGCGAGCTGATCGTGTGTTCTGCGCGCG 735
Db      462 GCGGAGCTGCGGAGACAGCGCGCGGTGGGGAAGGCCTTGGCGGGGTGCTCATCTGCTC 403
Qy      736 CTGGTCCGCGCGCGCTCGGCTTCTCTGTTCAACACCATATCCGCGCAGGTCTTTCATG 795
Db      402 ATCGCCGTGTACGTGGCCGGTTTCGGCTGGTCTTGGGGCCCCCTCGGGTGGCTGGTCCCC 343
Qy      796 GCGGACGTCCGCGCGCTGGCGTGGGCGCGCGCTGGGCAACATCGCGGTGATCGTGGCGC 855
Db      342 AGCGAGATCTTCCCGCTGGAGGTGCGCGCGCGCGCAGAGCGTCACGGTGGCGGTGAGC 283
Qy      856 CAGGAGATGCTGCTGTTTCATCATGGTGGGTGTTCTGTCATG 897
Db      282 TTCGCGTTACCGGTGTTCTGGCGCAGGCGTTCCTGTCTCATG 241

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 9, 2004, 13:30:53 ; Search time 524 Seconds  
(without alignments)  
9362.558 Million cell updates/sec

Title: US-10-089-787-1

Perfect score: 1083

Sequence: 1 atgctccctgctgctggccga.....ccacttgaagctgcttga 1083

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2941586 seqs, 226495651 residues

Total number of hits satisfying chosen parameters: 5883172

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA.\*

- 1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq.\*
- 2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq.\*
- 3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq.\*
- 4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq.\*
- 5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq.\*
- 6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq.\*
- 7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq.\*
- 8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq.\*
- 9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq.\*
- 10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq.\*
- 11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq.\*
- 12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq.\*
- 13: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq.\*
- 14: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq.\*
- 15: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq.\*
- 16: /cgn2\_6/ptodata/2/pubpna/US10C\_PUBCOMB.seq.\*
- 17: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq.\*
- 18: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq.\*
- 19: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq.\*
- 20: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1078.2	99.6	1083	9	US-09-815-242-7913
2	1078.2	99.6	1083	13	US-10-282-122A-30469
3	819.8	75.7	1080	13	US-10-282-122A-31717
4	767	70.8	1080	13	US-10-282-122A-33529
5	540.8	49.9	1080	13	US-10-282-122A-19657
6	539	49.8	1083	9	US-09-815-242-7594
7	537.6	49.6	1080	13	US-10-282-122A-23402
8	518.2	47.8	1083	13	US-10-282-122A-39002
9	518.2	47.8	1083	13	US-10-282-122A-39834
10	517.2	47.8	1083	9	US-09-815-242-9977
11	502.2	46.4	1083	9	US-09-741-669-258
12	502.2	46.4	1083	9	US-09-815-242-5931
13	502.2	46.4	1083	13	US-10-282-122A-20228
14	490.6	45.3	1060	13	US-10-282-122A-37049

15	480.4	44.4	1167	13	US-10-282-122A-15193
16	470.2	43.4	1083	13	US-10-282-122A-41275
17	460.6	42.5	1083	13	US-10-282-122A-41736
18	440.4	40.7	1167	13	US-10-282-122A-13829
19	412.6	38.1	1167	13	US-10-282-122A-12921
20	410	37.9	1083	13	US-10-282-122A-29880
21	409.2	37.8	1080	13	US-10-282-122A-28946
22	388.6	35.9	1083	13	US-10-282-122A-30712
23	368	34.0	1080	13	US-10-282-122A-32630
24	361.4	33.4	1083	9	US-09-815-242-7089
25	361.4	33.4	1083	13	US-10-282-122A-22235
26	361.4	33.4	1830121	15	US-10-329-960-1
27	361.4	33.4	1830121	16	US-10-329-960-1
28	326	30.1	645	13	US-10-282-122A-11904
29	253.4	23.4	1116	13	US-10-282-122A-8532
30	242.4	22.4	717	13	US-10-282-122A-25283
31	232.2	21.4	100848	13	US-10-672-787-39
32	231.4	21.4	1095	13	US-10-282-122A-36874
33	220.6	20.4	1071	15	US-10-156-761-6096
34	220.6	20.4	9025608	15	US-10-156-761-1
35	205.8	19.0	1077	13	US-10-282-122A-25879
36	175.4	16.2	1080	9	US-09-712-363-80
37	175.4	16.2	1080	13	US-10-282-122A-28470
38	174.2	16.1	640681	9	US-09-790-988-1
39	173.2	16.0	1082	13	US-10-335-977-836
40	171.8	15.9	1128	13	US-10-335-977-837
41	160.4	14.8	1062	9	US-09-815-242-7231
42	160.4	14.8	1062	13	US-10-282-122A-22534
43	153	14.1	570	13	US-10-335-977-835
44	148.2	13.7	1098	9	US-09-738-626-2368
45	148.2	13.7	3309400	9	US-09-738-626-1

ALIGNMENTS

RESULT 1

US-09-815-242-7913  
; Sequence 7913, Application US/09815242  
; GENERAL INFORMATION:  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Kari L.  
; APPLICANT: Zyskind, Judith W.  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John D.  
; APPLICANT: Carr, Grant J.  
; APPLICANT: Yamamoto, Robert T.  
; APPLICANT: Xu, H. Howard  
; TITLE OF INVENTION: Identification of Essential Genes in  
; FILE REFERENCE: ELITRA.011A  
; CURRENT APPLICATION NUMBER: US/09/815.242  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: 60/191,078  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: 60/206,848  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR APPLICATION NUMBER: 60/207,727  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: 60/242,578  
; PRIOR FILING DATE: 2000-10-23  
; PRIOR APPLICATION NUMBER: 60/253,625  
; PRIOR FILING DATE: 2000-11-27  
; PRIOR APPLICATION NUMBER: 60/257,931  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: 60/269,308  
; PRIOR FILING DATE: 2001-02-16  
; NUMBER OF SEQ ID NOS: 14110  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 7913  
; LENGTH: 1083  
; TYPE: DNA

```

961 ATGGCGCGATCCATCACCATTTTGGAACTGAAAGGCTGGCGGACCCGCGGTGATCGTG 1020
961 ATGGCGCGATCCATCACCATTTTGGAACTGAAAGGCTGGCGGACCCGCGGTGATCGTG 1020
1021 CGCTTCTGGATCATCAACGCTGATCCTGGTGTGATCGGCGCTCGCCACCTTGAAGCTGGT 1080
1021 CGCTTCTGGATCATCAACGCTGATCCTGGTGTGATCGGCGCTCGCCACCTTGAAGCTGGT 1080
1081 TGA 1083
1081 TGA 1083
1081 TGA 1083

RESULT 2
US-10-282-122A-30469
; Sequence 30469, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 30469
; LENGTH: 1083
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-10-282-122A-30469

Query Match 99.6%; Score 1078.2; DB 13; Length 1083;
Best Local Similarity 99.7%; Pred. No. 1.6e-270;
Matches 1080; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ATGCTCCTGCTGCTGCGCGCAATACCTGCAACAGTTCTTACAAGGCTTCGGCGTCTTCCAG 60
Db 1 ATGCTCCTGCTGCTGCGCGCAATACCTGCAACAGTTCTTACAAGGCTTCGGCGTCTTCCAG 60
QY 61 TAGCTGACCGCTGCGCGCAATTCACGCGTGCTCACGCGGTGTCGCTGTGCGTGTGCGCTG 120

```

Db 61 TACCTGACCTCGCGGCAATCTCAGCGTGTCAACGGCTGTGCTGCTGCTGCTG 120  
Qy 121 GGGCCCTGATGATCGTACCTTGCAGATCCCGGAGATCGCGGCGGCGGCGGCGGCGG 180  
Db 121 GGGCCCTGATGATCGTACCTTGCAGATCCCGGAGATCGCGGCGGCGGCGGCGGCGG 180  
Qy 181 GGTTCGCGAGTGCACCTGTGCGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 240  
Db 181 GGTTCGCGAGTGCACCTGTGCGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 240  
Qy 241 ACCGCATAGCATCAGCAGCTGTGCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 300  
Db 241 ACCGCATAGCATCAGCAGCTGTGCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 300  
Qy 301 GTGCTGGTGTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360  
Db 301 GTGCTGGTGTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360  
Qy 361 ATCGAAGAACTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420  
Db 361 ATCGAAGAACTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420  
Qy 421 GGCATCGGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480  
Db 421 GGCATCGGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480  
Qy 481 ATCGTCCCATGCTGAAGAGCTGCGAGTCCAGTTCGAGTTCGAGTTCGAGTTCGAGT 540  
Db 481 ATCGTCCCATGCTGAAGAGCTGCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGT 540  
Qy 541 TACTTCTCATGCTGCGGCTCGAGCAATGCGGTGAACCTCAGGAGGCTGCTGCTGCTG 600  
Db 541 TACTTCTCATGCTGCGGCTCGAGCAATGCGGTGAACCTCAGGAGGCTGCTGCTGCTG 600  
Qy 601 GGCATCATGCGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660  
Db 601 GGCATCATGCGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660  
Qy 661 AAGTGAAGTTCGCGAGTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720  
Db 661 AAGTGAAGTTCGCGAGTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720  
Qy 721 GTGCTTCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780  
Db 721 GTGCTTCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780  
Qy 781 GCGCAGCTCTTCATGCGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840  
Db 781 GCGCAGCTCTTCATGCGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840  
Qy 841 GCGGTGATGCTGCGGAGGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900  
Db 841 GCGGTGATGCTGCGGAGGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900  
Qy 901 ACCCTCTCGGTGATGATCCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960  
Db 901 ACCCTCTCGGTGATGATCCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960  
Qy 961 ATGGCGCGATCCATCAACCTTTCGAAAGGCTGCGGAGGAGGAGGAGGAGGAGGAGG 1020  
Db 961 ATGGCGCGATCCATCAACCTTTCGAAAGGCTGCGGAGGAGGAGGAGGAGGAGGAGG 1020  
Qy 1021 CGCTTCTGGATCATACCGGTGATCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1080  
Db 1021 CGCTTCTGGATCATACCGGTGATCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080  
Qy 1081 TGA 1083  
Db 1081 TGA 1083

RESULT 3

US-10-282-122A-31717  
; Sequence 31717, Application US/10282122A  
; Publication No. US20040029129A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Liangsu  
; APPLICANT: Zamudio, Carlos  
; APPLICANT: Malone, Cheryl  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Kari  
; APPLICANT: Zvakind, Judith  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John  
; APPLICANT: Carr, Grant  
; APPLICANT: Yamamoto, Robert  
; APPLICANT: Forsyth, R.  
; APPLICANT: Xu, H.  
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
; FILE REFERENCE: ELITRA 034A  
; CURRENT APPLICATION NUMBER: US/10/282,122A  
; CURRENT FILING DATE: 2003-02-20  
; PRIOR APPLICATION NUMBER: 60/191,078  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: 60/206,848  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR APPLICATION NUMBER: 60/207,727  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: 60/230,335  
; PRIOR FILING DATE: 2000-09-06  
; PRIOR APPLICATION NUMBER: 60/230,347  
; PRIOR FILING DATE: 2000-09-09  
; PRIOR APPLICATION NUMBER: 60/242,578  
; PRIOR FILING DATE: 2000-10-23  
; PRIOR APPLICATION NUMBER: 60/253,625  
; PRIOR FILING DATE: 2000-11-27  
; PRIOR APPLICATION NUMBER: 60/257,931  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: 60/267,636  
; PRIOR FILING DATE: 2001-02-09  
; PRIOR APPLICATION NUMBER: 60/269,308  
; PRIOR FILING DATE: 2001-02-16  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 78614  
; SOFTWARE: Patent in version 3.1  
; SEQ ID NO 31717  
; LENGTH: 1080  
; TYPE: DNA  
; ORGANISM: Pseudomonas putida  
US-10-282-122A-31717

Query Match Similarity 75.7%; Score 819.8; DB 13; Length 1080;  
Best Local Similarity 85.0%; Pred. No. 2.3e-203;  
Matches 917; Conservative 0; Mismatches 162; Indels 0; Gaps 0;

Qy 1 ATGCTCTGCTGCTGCGGAGTACCTGCAACAGTTCCTCAAGGCTTCGCGCTCTTCAG 60  
Db 1 ATGCTCTGCTGCTGCGGAGTATCTGCAACAGTTCCTCAAGGCTTCGCGCTCTTCAG 60  
Qy 61 TACCTGACCTCGCGGCAATTCCTCAGCGTGTCTACCGGCTGTGCTGTGCTGTGCGTG 120  
Db 61 TACCTGTCCCTCGCGGAGTTCCTTGGTGTACTGACCGGCTGTGCTGTGCGCTGTGCGTG 120  
Qy 121 GGGCCCTGATGATCGTACCTTGCAGATCCCGGAGATCCCGGAGGCGGCGGCGGAGCAG 180  
Db 121 GGGCCCTGATGATCGTACCTTGCAGATCCCGGAGATCCCGGAGGCGGCGGCGGAGCAG 180  
Qy 181 GGTTCGCGAGTGCACCTGTGCGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 240  
Db 181 GGTTCGCGAGTGCACCTGTGCGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 240  
Qy 241 ACCGCATAGCATCAGCAGCTGTGCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 300  
Db 241 TCGGCGATCGCGGTGAGCAGCCTTGTGTGGGCGGAGGAGGAGGAGGAGGAGGAGG 300

301 GTGCTGGTCGTTACCTGCTGTTCCGTCGCCATCGGCTGGGTAGACGACTACCGCAAGGTG 360  
Db GTGCTGATGTCACCTCGGCTTCGGTGGCAATCGCTGGGTGATGACTACCGCAAGGTG 360  
361 ATCGAAGAACTCCGTCGGCTCGCGAGCGCTGGAACTCTTCTGGCAGTGGGTTC 420  
Db ATCGAAGAACTCCGTCGGCTTCGAGAGCGCTGGAACTCTTCTGGCAGTGGGTTC 420  
421 GGCAATCGCGCGCGCTGTTCTCTACATGACTCGCGAAACCCCGATCGAGACCAACCTG 480  
Db GGCTCGCGCGCGCTGTTCTCTACAGACGGCGCAACACGCTCGAGACCAACCTG 480  
481 ATGCTGCGGATGCTGAAGAGCGTGAATCCAGTTGGGCATCTTCTTCTGGTGGTCTGACC 540  
Db ATCTGCGGTTTCATCAGGATGTACCAATTCGTTGGCGCTCGGCTTCGTCGTAAGACC 540  
541 TACTTCGTCATCGTCGGCTCGAGCAATCGCGTGAACCTCACGACGCTCTCGACGGCTG 600  
Db TACTTCGTCATCGTCGGCTCGAGCAACCGCGTCAATCTCACGATGGCTTGAGCGCTG 600  
601 CGGATCATCGGACGGTAATGGTTCCCGCGCGCTGGGCATCTTCTGCTACCTGTCGGC 660  
Db GCAATCATCGGACGGTATGTCGCGCGCGCTGGGCATCTTCTGCTACCTGTCGGT 660  
661 AACGTGAAGTTCCGCGAGTACTGCTGATTCGCAACGTACCGGCGCGCGGAGCTGATC 720  
Db AACGTGAAGTTCCGCGAATACCTGCTGATCCCTACGTGCGCGGCTCGGCGAATGATC 720  
721 GTGTTCTCGCGCGCTGTCGGCGCGCTCGGCTTCCTCTGCTTCAACACCTATCCG 780  
Db GTGTTCTCGCGCGCTGATCGTTCGCGCGCTGGCTTCTGTTCAACACCTATCCG 780  
781 GCGAGGTTCTCATGGGCGAGTCCGCGCGCTGGCGCTGGCGCGCGCTGGGACCATC 840  
Db GCGAAGTTCTCATGGGCGAGTCCGCGCGCTGGCGCTGGCGCGCGCTGGGACCATC 840  
841 GCGGTGATGTCGCGCGAGAGATCGTGTTCATCATGGGTGGGTGCTTCTGTCATGGAA 900  
Db GCGGTGATGTCGCGCGAGAGATCGTGTTCATCATGGGTGGGTGCTTCTGTCATGGAA 900  
901 ACCCTCTGGTATGATCAGGTGCTTCTTCAAGCTGACCGAGCGCGCTTCTCGT 960  
Db ACCCTCTGGTATGATCAGGTGCTTCTTCAAGCTGACCGAGCGCGCTTCTCGC 960  
961 ATGGCGCGATCATCACCATTGCAATCTGAAGCTGGCGCGCGCGCGCTGATCGT 1020  
Db ATGGCGCGATCATCACCATTGCAATCTGAAGCTGGCGCGCGCGCGCTGATCGT 1020  
1021 CGCTTCTGATCATCAGGTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1079  
Db CGTTCCTGATCATCAGGTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1079

RESULT 4

US-10-282-122A-33529  
; Sequence 33529, Application US/10282122A  
; Publication No. US20040029129A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Liangsu  
; APPLICANT: Zamudio, Carlos  
; APPLICANT: Malone, Cheryl  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Karl  
; APPLICANT: Zyskind, Judith  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John  
; APPLICANT: Carr, Grant  
; APPLICANT: Yamamoto, Robert  
; APPLICANT: Forsyth, R.  
; APPLICANT: Xu, H.  
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
; FILE REFERENCE: ELITRA.034A  
; CURRENT APPLICATION NUMBER: US/10/282,122A

CURRENT FILING DATE: 2003-02-20  
; PRIOR APPLICATION NUMBER: 60/191,078  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: 60/206,848  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR APPLICATION NUMBER: 60/207,727  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: 60/230,335  
; PRIOR FILING DATE: 2000-09-06  
; PRIOR APPLICATION NUMBER: 60/230,347  
; PRIOR FILING DATE: 2000-09-09  
; PRIOR APPLICATION NUMBER: 60/242,578  
; PRIOR FILING DATE: 2000-10-23  
; PRIOR APPLICATION NUMBER: 60/253,625  
; PRIOR FILING DATE: 2000-11-27  
; PRIOR APPLICATION NUMBER: 60/257,931  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: 60/267,636  
; PRIOR FILING DATE: 2001-02-09  
; PRIOR APPLICATION NUMBER: 60/269,308  
; PRIOR FILING DATE: 2001-02-16  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 78614  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 33529  
; LENGTH: 1080  
; TYPE: DNA  
; ORGANISM: Pseudomonas syringae  
; US-10-282-122A-33529

Query Match 70.8%; Score 767; DB 13; Length 1080;  
Best Local Similarity 81.9%; Pred. No. 1.2e-189;  
Matches 884; Conservative 0; Mismatches 195; Indels 0; Gaps 0;

Qy 1 ATGCTCTGCTGCTGCGCGAATACCTGCAACAGTTCTTACAAAGGCTTTTCCGTTCTTCAG 60  
Db 1 ATGCTCTGCTGCTGCGCGAATACCTGCAACAGTTCTTACAAAGGCTTTTCCGTTCTTCAG 60  
Qy 61 TACCTGACCTGCGCGGCAATCTCAGCGTGTCTCAGCGCTGTCTGCGTGTGCGTGTG 120  
Db 61 TACCTGCTCTGCGCGGATCTTCTGCGTGTCTCAGCGTGTCTGCGTGTGCGTGTG 120  
Qy 121 GGCGCTGATGATCGGTACTTCTGAGATCCCGAGATCGCGAGTGGCGAGTGGCGACGAC 180  
Db 121 GGCGCGTGTGATCGCGACCTCTGCGAGATGCGCGAGTGGCGAGTGGCGTTCGTAACGAT 180  
Qy 181 GGTCCGCGAGTCCGACCTGTGGAAGAGGCGACCCCGACCATGGCGCGCGCTGATCCTT 240  
Db 181 GGTCCGCGATCGGACCTGTGTCAGTCCGATCGCGACCATGGCGCGCGCTGATCCTC 240  
Qy 241 ACCGCCATAGCCATCAGCAGCGTGTGTTGGCGGATCTTTCCAAACCGCTACGTGGGTA 300  
Db 241 TCGTCCATCGGTATCAGCACCTTCTGCTGGGCTGATTTGAGCAACCGTTACGTCGGGT 300  
Qy 301 GTGCTGTGTTACCTGCTGTTTCGGTGCATCGGCTGGGTAGACGACTACCGCAAGGTG 360  
Db 301 GTGTTGCTGTGACCTTCTGTTTCGGTGCATCGGCTGGGTAGTACTACCGCAAGGTG 360  
Qy 361 ATCGAAGAACTCCGTCGGCTCGCGAGCGCTGGAAGTACTTCTTGGCAGTGGGTGTTTC 420  
Db 361 ATTGAAAGAAATCCCGTGTGCTGCGAGTGGTGGAGTATTTCTTGGCAGTGGGTGTTTC 420  
Qy 421 GGATCGCGCGCGCTGTTCTTCTACATGACTCGCGAAACCCCGATCGAGACCAACCTG 480  
Db 421 GGCTTTGCGCGGCAATCTTCTTGTACAGACTGCGCCATCGCGACCAACCAATG 480  
Qy 481 ATCGTCCGATGCTGAAGAGCGTTCGAGATCCAGTTGGGATCTTCTTGGTGTCTGACC 540  
Db 481 ATCGTCCGATGCTGAAGAGCGTTCGAGATCCAGTTGGGATCTTCTTGGTGTCTGACC 540  
Qy 541 TACTTCGTCATCGTCGGCTCGAGCAATGGGTGAACCTCACGACGCTCTCGACGGCTG 600  
Db 541 TACTTTGTGATGTCGGTTCACGCAACGAGTCAACCTGACTACGGCTTGGAGCGGCTG 600



Db 841 GCCGTACTGCTGCTCAGGAGTTCTCTGCTGGTGTATCAATGGCGCGCTGTTTGTGGTTGAA 900  
 Qy 901 ACCCTCTCGGTGATGATCCAGGTCGCTTCCCTTCAAGCTGACCGGACCGCGCTTCCGT 960  
 Db 901 ACCCTCTCGGTGATTTCCAGGTCGCTTCCCTTCAAGCTGACCGGACCGCTTCCGT 960  
 Qy 961 ATGGCGCGATCCATCACCATTTCGAACTGAAAGGCTGGCGGACCGCGGCTGATCGTG 1020  
 Db 961 ATGGCGCGATTCACCACTATGAACTGAAAGGCTGGCGGACCGCGGCTGATCGTG 1020  
 Qy 1021 CGCTTCTGGATCATCACCGTATCTGCTGCTGATCGGCTCGCCACCTTTGAAGCTCGT 1080  
 Db 1021 CGCTTCTGGATTAATTCGCTGATGCTGCTGATTTGGCCTGGCAACGCTGAGGTACGT 1080

RESULT 6

US-09-815-242-7594  
 ; Sequence 7594, Application US/09815242  
 ; Patent No. US2002061569A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Haselbeck, Robert  
 ; APPLICANT: Ohlsen, Karl L.  
 ; APPLICANT: Zyskind, Judith W.  
 ; APPLICANT: Wall, Daniel  
 ; APPLICANT: Trawick, John D.  
 ; APPLICANT: Carr, Grant J.  
 ; APPLICANT: Yamamoto, Robert T.  
 ; APPLICANT: Xu, H. Howard  
 ; TITLE OF INVENTION: Identification of Essential Genes in  
 ; FILE OF INVENTION: Prokaryotes  
 ; FILE REFERENCE: ELITRA.011A  
 ; CURRENT APPLICATION NUMBER: US/09/815,242  
 ; CURRENT FILING DATE: 2001-03-21  
 ; PRIOR APPLICATION NUMBER: 60/191,078  
 ; PRIOR FILING DATE: 2000-03-21  
 ; PRIOR APPLICATION NUMBER: 60/206,848  
 ; PRIOR FILING DATE: 2000-05-23  
 ; PRIOR APPLICATION NUMBER: 60/207,727  
 ; PRIOR FILING DATE: 2000-05-26  
 ; PRIOR APPLICATION NUMBER: 60/242,578  
 ; PRIOR FILING DATE: 2000-10-23  
 ; PRIOR APPLICATION NUMBER: 60/253,625  
 ; PRIOR FILING DATE: 2000-11-27  
 ; PRIOR APPLICATION NUMBER: 60/257,931  
 ; PRIOR FILING DATE: 2000-12-22  
 ; PRIOR APPLICATION NUMBER: 60/269,308  
 ; PRIOR FILING DATE: 2001-02-16  
 ; NUMBER OF SEQ ID NOS: 14110  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 7594  
 ; LENGTH: 1083  
 ; TYPE: DNA  
 ; ORGANISM: Klebsiella pneumoniae  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: (1)...(1083)  
 ; US-09-815-242-7594

Query Match 49.8%; Score 539; DB 9; Length 1083;  
 Best Local Similarity 68.6%; Pred. No. 2.3e-130;  
 Matches 743; Conservative 0; Mismatches 340; Indels 0; Gaps 0;

Qy 1 ATGCTCTGCTGCTGCGGATACCTGCAAGTTCATACAGGCTTCGGCGCTTCCAG 60  
 Db 1 ATGTTAGTAGTGGCTGCGGCAACACTTGGTCAAAATATATTCGGCTTTAACGCTTTTCA 60  
 Qy 61 TACCTGACCTGCGGCGGATCTTCAGCGTGCTCACCGGCTGTCGCTGTCGTGGCTG 120  
 Db 61 TATCTGACGTTTCGCGGATCGTGCAGCTGCTGACCGGCTGTTTCATCTCGTTGGATG 120  
 Qy 121 GGGCGCTGGATGATCGGTACCTTGGAGATCCCCAGATCGGCGAGCGCTGGCGAACGAC 180  
 Db 121 GGGCGCGCATGATCGCCGCTCTGCAAAAACCTCGCCTTTGGCGAGGCTGCTGTAACGAC 180

Qy 181 GGTCCGAGTGCACCTGTCTGAAGAGGGGACCCCGACCATGGGGGGCGCCTGTATCTT 240  
 Db 181 GGGCGGAGTCTCATTTTCAGTAAACCGGTATCTCGACCATGGGGGGGATCATGATCTC 240  
 Qy 241 ACCGCATAGCATCAGCACGCTGTGTGGGCGATCTTTCCAAACCGCTAGCTGTGGGTA 300  
 Db 241 ACCGGATCACGTTTCGTTCTGTGTGGGCTATCCATCTAACCCGCTACGTCTGGTGC 300  
 Qy 301 GTGCTGTGTTTACCCCTGCTGTTCGCTGCTATCGCTGGGTAGACGATACCGCAAGTG 360  
 Db 301 GTACTGACGGTATTAAATCGGCTACCGCATCATCGGTTTCGTTGATGATTACCGTAAAGTC 360  
 Qy 361 ATCGAAGAACTCCCGTGGCTGCGGAGCGCTGGAAGTACTTCTGGCAGTCGCGTTC 420  
 Db 361 GTGGCAAGATACCAAGGCTGATCGCCGCTGGAAGTATTTCTGATGTGGTATC 420  
 Qy 421 GGCATCGGCGCGCGTGTTCCTTACATGATGCTGCCGAAACCCCGATCGAGACACCTG 480  
 Db 421 GGCCTGGGCGTGGCCTTCGCGCTGTATCTGGCGGGAAGATACCCCGCAACCGAGCTG 480  
 Qy 481 ATCGTCCGATGCTGAAGAGCGTCGAGATCCAGTTGGGATCTTCTCTGTGCTCTGACC 540  
 Db 481 GTGGTCCGTTCTTTAAAGACGTAATGCCGAGCTGGGGCTGTTCTATATCTTGTGGCC 540  
 Qy 541 TACTTCTCATGCTCGGCTCGAGCAATCGGTGAACCTCACCGACGCTTCGACGCGCTG 600  
 Db 541 TACTTCTGTTATCGTGGTACCGGCAACCGCTCAACCTGACCGACGCGCTCGACGCGCTG 600  
 Qy 601 GGCATCGGCGCGGTAATGTTGCGGCGCGCTGGGATCTTCTGTACCTGCTGCGGC 660  
 Db 601 GCGATTATCGCACCGTTTCTGCGCGGAGGCTTGGCGTGGTGGCATGCGGCCACCGC 660  
 Qy 661 AACGTGAAGTTCGCCGAGTACCTGCTGATTCCCAAGTACCGGCGCGCGCGAGCTGATC 720  
 Db 661 AACATGAATTCGCCCAACTACCTGCATATCCGTAATCTCGCCACGCTGGCGAGCTG 720  
 Qy 721 GTGTTCTGCGCGCGCTGCTGCGCGCGGCTCGGCTTCTCTGTGTTCAACACCTATCCG 780  
 Db 721 ATGCTCTGACGCGGATTCGCGGCGGCGCTGGGCTTCTGTGTTCAACACCTATCCCG 780  
 Qy 781 GCGCAGGTCTTCATGGCGGACGTCGCGCGGCTGGGCGCGCGCGCTGGGCGCACCATC 840  
 Db 781 GCGCAAGTCTTTATGGCGGACGTCGCTTCACTGGCGCTCGCGGCGGCTGGCATTATC 840  
 Qy 841 GCGGTGATGCTGCGCGGAGAGTCTGCTGTTTCATCATGCTGGTGGGTGTTCTCATGGA 900  
 Db 841 GCGGTGCTGCTGCGTCAAGAGTTCCTGCTGCTGATCATGCGCGGCGGCTTTTCGTTGGTGA 900  
 Qy 901 ACCCTCTCGGTGATGATCCAGGTCGCTTCTTCAAGTGAACGCGCGCGCTTTCGCT 960  
 Db 901 ACCTGTCGGTATTCGAGGTCGCTTCTTAACTGCGGCTCAGGCGCATCTTCCG 960  
 Qy 961 ATGGCGCGGATCCATCACCATTTCGAACTGAAAGGCTGGCGGACCGCGCGCTGATCGTG 1020  
 Db 961 ATGGCGCGGATCCACCACTATGAACTGAAAGGCTGGCGGAGCGCGGCTTATCGTG 1020  
 Qy 1021 CGCTTCTGGATCATCCCGTATCTGCTGCTGATCGGCTCGCCACCTTTGAAGCTGCGT 1080  
 Db 1021 CGCTTCTGGATTAATTCGCTGATGCTGCTGCTGATTTGGCCTGGCAACGCTGAGGTACGT 1080  
 Qy 1081 TGA 1083  
 Db 1081 TAA 1083

RESULT 7

US-10-282-122A-23402  
 ; Sequence 23402, Application US/10282122A  
 ; Publication No. US2004029129A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Wang, Liangsu  
 ; APPLICANT: Zamudio, Carlos

APPLICANT: Malone, Cheryl  
 APPLICANT: Haselbeck, Robert  
 APPLICANT: Ohlsen, Kari  
 APPLICANT: Zykkind, Judith  
 APPLICANT: Wall, Daniel  
 APPLICANT: Trawick, John  
 APPLICANT: Carr, Grant  
 APPLICANT: Yamamoto, Robert  
 APPLICANT: Forsyth, R.  
 APPLICANT: Xu, H.  
 TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
 FILE REFERENCE: ELITRA.034A  
 CURRENT APPLICATION NUMBER: US/10/282,122A  
 CURRENT FILING DATE: 2003-02-20  
 PRIOR APPLICATION NUMBER: 60/191,078  
 PRIOR FILING DATE: 2000-03-21  
 PRIOR APPLICATION NUMBER: 60/206,848  
 PRIOR FILING DATE: 2000-05-23  
 PRIOR APPLICATION NUMBER: 60/207,727  
 PRIOR FILING DATE: 2000-05-26  
 PRIOR APPLICATION NUMBER: 60/230,335  
 PRIOR FILING DATE: 2000-09-06  
 PRIOR APPLICATION NUMBER: 60/230,347  
 PRIOR FILING DATE: 2000-09-09  
 PRIOR APPLICATION NUMBER: 60/242,578  
 PRIOR FILING DATE: 2000-10-23  
 PRIOR APPLICATION NUMBER: 60/253,625  
 PRIOR FILING DATE: 2000-11-27  
 PRIOR APPLICATION NUMBER: 60/257,931  
 PRIOR FILING DATE: 2000-12-22  
 PRIOR APPLICATION NUMBER: 60/267,636  
 PRIOR FILING DATE: 2001-02-09  
 PRIOR APPLICATION NUMBER: 60/269,308  
 PRIOR FILING DATE: 2001-02-16  
 Remaining Prior Application data removed - See File Wrapper or PALM.  
 NUMBER OF SEQ ID NOS: 78614  
 SOFTWARE: PatentIn version 3.1  
 SEQ ID NO 23402  
 LENGTH: 1080  
 TYPE: DNA  
 ORGANISM: Klebsiella pneumoniae  
 US-10-282-122A-23402

Query Match 49.6%; Score 537.6; DB 13; Length 1080;  
 Best Local Similarity 68.6%; Pred. No. 5.4e-130;  
 Matches 741; Conservative 0; Mismatches 339; Indels 0; Gaps 0;

QY 1 ATGCTCCTGCTGCTGCGGCAATACCTGCAACAGTCTTCAAGGGCTTCGGCGTCTTCAG 60  
 DB 1 ATGTTAGTATGCTGCGGCAACACTTGTGTCAAATATTATTCGGGCTTTAAAGCTCTTTTCA 60  
 QY 61 TACCTGAACCTGCGGGCAATCTCAGCGTGTCAAGGCTGCTGCTGCTGCTGCTGCTGCTG 120  
 DB 61 TATCTGACGTTTTCGGGCCATCTGTCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120  
 QY 121 GGGCCCTGATGATCGGTACCTTGAGATCCCGGATGCTGCTGCTGCTGCTGCTGCTGCTG 180  
 DB 121 GGGCCCGCATGATCGCGCTGCTGCAAAACTCGCCTTTGGCCAGTCTGCTGCTGCTGCTG 180  
 QY 181 GTTCGCACTCGCACTGCTGCAAGAGGCGCCCGGATGCTGCTGCTGCTGCTGCTGCTGCTG 240  
 DB 181 GGGCCCGAGTCTCATTTTCAGTAAACCGGTATCTCGGACCATGCGGGGATCATGATCTCT 240  
 QY 241 ACCGCCATAGCCATCAGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300  
 DB 241 ACCGGATCAGCGTTTCGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300  
 QY 301 GTGCTGCTGCTTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360  
 DB 301 GTACTACCGGTATTAATTCGGCTACGCACTCATCGTTCGTTGATGATTAACCGTAAAGTC 360  
 QY 361 ATCAGAAGAACTCCCGTGGCTGCGGAGTACTTCTGCGAGTCGCTGCTGCTGCTGCTGCT 420

361 GTGCGCAAGATACCAAGGCCTGATCGCCGCTGGAAGTATTCTTGGATGTCGGTGATC 420  
 QY 421 GGCATCGGGCGCGCGCTGTTCTCTACATGACATGCGGAACCCCGATCGAGACACCCCTG 480  
 DB 421 GCTTGGCGTGGCTTTCGGCTGATCTGCGGGGAAAGATACCCCGGGAACCGAGCTG 480  
 QY 481 ATCGTCCGATGCTGAAGAGCGTCGAGATCCAGTTGGGCACTCTTCTTGGTCTTCGACC 540  
 DB 481 GTGGTCCGTTCTTTAAAGACGTAATCCCGCAGCTGGGGCTGTTCTATATCTTCTTGGCC 540  
 QY 541 TACTTGTCTATCGTCCGCTCGAGCAATGGGTGAACCTCACCGACGCTCTCGACGCGCTG 600  
 DB 541 TACTTGTCTATCGTCCGCTACCGCAACGCGTCACTGACCGACGCTCTCGACGCGCTG 600  
 QY 601 GGCATCATCGGACGCGTAATGTTGCGCGCGCTCGGCACTCTTCTGCTACCTCTGCGGC 660  
 DB 601 GCGATTATGCGGACCGTTTTCGTCGCGGCGAGGCTTTCGCTGCTGCTGCTGCTGCTGCTG 660  
 QY 661 AACGTGAAGTTCCGCGAGTACCTGCTGATTCCTCAAGTACCGGCGCGCGCGGAGCTGATC 720  
 DB 661 AACATGAATTTCCGCAACTACCTACATATCCCGTATCTGCGCCACCGCGCGAGCTGGT 720  
 QY 721 GTGTTCTGCGCGCGCTGCTGCGCGCGCGCTCGGCTTCTCTGTTCAACACTATCCG 780  
 DB 721 ATCGTCTGTACGCGGATGTTGCGGGCGGCTGCGGCTTCTGTTGTTCAACACTATCCG 780  
 QY 781 GCGAGGTTCTTATGCGGACGTCGCGCGCGCTGCGGCTGCGGCGCGCGCGCGGCTGCGAC 840  
 DB 781 GCGCAAGTCTTTATGCGGCGAGCTGCGGTTCACTGCGCGCTGCGCGCGCGCTTTCGCA 840  
 QY 841 GCGGTGATCGTGGCGGAGATCGTGTCTTCTATCATGGGTGGGCTGTTCTGCTCATGGAA 900  
 DB 841 GCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900  
 QY 901 ACCTCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 960  
 DB 901 ACCTGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960  
 QY 961 ATGGCGCGGATCCATCACCATTTCGAAGTGAAGGCTGGCGGCGCGCGCGCGCTGATCG 1020  
 DB 961 ATGGCGCGGATCCACACCACTATGAAGTGAAGGCTGGCGGCGCGCGCGCGCTGATCG 1020  
 QY 1021 CGTCTTGGATCATCACCGTATCGTGGTCTGATCGGCTCGGCTCGGCTCGGCTCGGCT 1080  
 DB 1021 CGTCTTGGATTTATTTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1080

RESULT 8  
 US-10-282-122A-39002  
 ; Sequence 39002, Application US/10282122A  
 ; Publication No. US20040029129A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Wang, Liangsu  
 ; APPLICANT: Zamudio, Carlos  
 ; APPLICANT: Malone, Cheryl  
 ; APPLICANT: Haselbeck, Robert  
 ; APPLICANT: Ohlsen, Kari  
 ; APPLICANT: Zykkind, Judith  
 ; APPLICANT: Wall, Daniel  
 ; APPLICANT: Trawick, John  
 ; APPLICANT: Carr, Grant  
 ; APPLICANT: Yamamoto, Robert  
 ; APPLICANT: Forsyth, R.  
 ; APPLICANT: Xu, H.  
 ; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
 ; FILE REFERENCE: ELITRA.034A  
 ; CURRENT APPLICATION NUMBER: US/10/282,122A  
 ; CURRENT FILING DATE: 2003-02-20  
 ; PRIOR APPLICATION NUMBER: 60/191,078  
 ; PRIOR FILING DATE: 2000-03-21  
 ; PRIOR APPLICATION NUMBER: 60/206,848  
 ; PRIOR FILING DATE: 2000-05-23  
 ; PRIOR APPLICATION NUMBER: 60/207,727

```
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 39002
; TYPE: DNA
; LENGTH: 1083
; ORGANISM: Salmonella typhimurium
US-10-282-122A-39002

Query Match      47.8%; Score 518.2; DB 13; Length 1083;
Best Local Similarity 67.4%; Pred. No. 5.9e-125;
Matches 730; Conservative 0; Mismatches 353; Indels 0; Gaps 0;

QY      1  ATGTCCTGCTGCTGGCGCAATACCTGCAACAGTCTACAGGGCTTCGGCGTCTTCCAG 60
DB      1  ATGTTAGTTTGGCTGGCGAGCATTTGGTCAATATATATCCGGCTTTAAGCTTTTCT 60

QY      61  TACGTGACCTTGGCGGCAATCTCAGGCTGTCTACCGGCTGTCCGCTGTGCTGCTG 120
DB      61  TATCTGACGTTTGGCGGCAATCTCAGGCTGTCTACCGGCTGTCTCCTTTATGATG 120

QY      121  GGGCCCTGGATGATCCGTACCTTGCAGATCCCGAGATCGGCCAGCGGCTGCGCAACGAC 180
DB      121  GCGCCGCTATGATCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180

QY      181  GGTCCGAGTGGCACTCTCGAAGAGGGCAACCGGACCGGCTGTGCTGCTGCTGCTGCTG 240
DB      181  GGGCCGGAATGGCACTTCAGTAAACCGGCTGCGGACCGGCTGCTGCTGCTGCTGCTG 240

QY      241  ACCGCTATGCAATCAGACGCTGTGCGGCGATCTTCCACCGCTAGCTGCTGGTA 300
DB      241  ACGTCGATGATGATTCGGCTTCTGTTATGCGGCTTACCGCTTACCGGCTAGCTGCTG 300

QY      301  GTGCTGGTCTGCTTACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360
DB      301  GTGCTGGTGGTATGATCGGCTAGCGCATTTATCGGTTTTTGTGATGATGATCGCAAGT 360

QY      361  ATCGAGAAGAACTCCGCTGGCTGCGGAGCGCTGGAAGTACTTTCGCGAGTCCGTTTC 420
DB      361  GTGCGGAAGATACCAAAAGGCTGATTCGCGCTGGAATATTTCTGGATGCTGGTTATC 420

QY      421  GGCATCGGCGCGCGCTGCTCTCTACATGACTGCGCAAAACCGGATCGAGACCACTG 480
DB      421  GCGCTCGGCTGGCTTTTGGCTTTATCTGCTGGGAAGACACGCGCGGACCACTG 480

QY      481  ATCGTGGCAATGCTGAAGAGCTGCAATGATGATGATGATGATGATGATGATGATGATG 540
DB      481  GTGCTGGCGCTTTTAAAGATGTTATGCGCAATTTGGGCTGTTTACATTTCTGCTGCT 540

QY      541  TACTTCTGATGCTGCGCTCGAGCAATGCGTGAACCTCACCGAGCTCTCGAGCGCTG 600
DB      541  TACTTTGATGCTGCGTACGGTAAACCGCTGAACCTGACCGAGCGGCTGATGATGCTG 600

QY      601  GCGATCATGCGCAAGGTAATGTTCCCGGCGGCTGCGGCACTTCTGCTGCTGCTGCTG 660
DB      601  GCGATATGCGCAAGTCTTCTGCTGCGGCGGCTTGGCTGCTGCTGCTGCTGCTGCTG 660

QY      661  AAGCTGAAGTTCCGCGGATACCTGCTGATCCCAAGTACCGGCGCGCGGAGCTGATC 720
```

## RESULT 9

```
US-10-282-122A-39834
; Sequence 39834, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
```

```
DB      661  AACATGAATTCGCAATTAATCTGATATTCGATATTTACGCCATGCGGCGAGCTG 720
QY      721  GTGTTCTGCGCGCGCTGCTGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780
DB      721  ATTGTCTGTACGGCGAATTTGTCGGCGGGGATTTAGGATTTCTTTGTTTAAACCTATCCG 780
QY      781  GCGCAGGCTTTCATGGCGGAGTCCGCGCGCTGCGCTGCGCTGCGCGCGCTGCGCGC 840
DB      781  GCGCAGGCTTTTATGGCGGAGTCCGATCGCTGCGCTGCGCTGCGCTGCGCTGCGCT 840
QY      841  GCGGTGATGCTGCGCGCGGAGATCGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
DB      841  GCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
QY      901  ACCCTCTGCTGATGATCCAGTCTGCTTCTTCAAGCTGACCGGACCGCGCTTCCGT 960
DB      901  ACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960
QY      961  ATGGCGCGATCCATCAGATTTTCAAGCTGAGAGGCTGCGCGGCGCGCTGCTGCTG 1020
DB      961  ATGGCGCTTATCCATCAGCACTATGAACTGAAAGGCTGCGCGGCGCGCTGCTGCTG 1020
QY      1021  CGCTTCTGATCATCACCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080
DB      1021  CGCTTCTGATTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080
QY      1081  TGA 1083
DB      1081  TAA 1083
```

; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 78614  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 39834  
; LENGTH: 1083  
; TYPE: DNA  
; ORGANISM: Salmonella typhi  
US-10-282-122A-39834

Query Match 47.8%; Score 518.2; DB 13; Length 1083;  
Best Local Similarity 67.4%; Pred. No. 5.9e-125;  
Matches 730; Conservative 0; Mismatches 353; Indels 0; Gaps 0;

QY	1	ATGCTCTGCTGCTGCGGCAATCTTCAAGAGTCTTCAAGGGCTTCGGCTCTTCCAG	60
DB	1	ATGTTAGTTGGCTGCGCGAGCAATTTGGTCAATATATTCCGGCTTTAAACGCTCTTTCT	60
QY	61	TACTGACCTGCGCGGCAATCTTCAAGAGTCTTCAAGGGCTTCGGCTCTTCCAG	120
DB	61	TATCTGACGTTTGGCGCCATCGTCAAGCTTTGACCGGCTGTTTCATCTCTTTATGATG	120
QY	121	GGGCTCTGATGATCGGTACCTTGGTACGATCCCGGAGTGGCGGCGTGGCGACGAC	180
DB	121	GGGCTCTGATGATCGGTACCTTGGTACGATCCCGGAGTGGCGGCGTGGCGACGAT	180
QY	181	GGTCCGAGTCGACCTGTCGAAGAAAGGCGACCCGACCATGGCGGCGGCGTGTCTT	240
DB	181	GGGCGGAGTCGACCTTCAAGAAAGGCGACCCGACCATGGCGGCGGCGTGTCTT	240
QY	241	ACGCGATAGCATGACGAGTCTGTTGGGCGGATCTTTTCCAAACGCTACGTGGGTA	300
DB	241	ACGCGATAGCATGACGAGTCTGTTGGGCGGATCTTTTCCAAACGCTACGTGGGTA	300
QY	301	GTCTGCTGCTTACCTGCTGTTGCGTGCATCGGCTGGGTACGACACTACCGAAGTG	360
DB	301	GTCTGCTGCTTACCTGCTGTTGCGTGCATCGGCTGGGTACGACACTACCGAAGTG	360
QY	361	ATCGAAGAACTCCGTCGCTGCGGAGCGGCTGGAAGTACTTTCGAGTGGGTTC	420
DB	361	ATCGAAGAACTCCGTCGCTGCGGAGCGGCTGGAAGTACTTTCGAGTGGGTTC	420
QY	421	GGCATGCGGCGCGGCTGTTCTCATAGTACGTCGCGAAGCGGATCGAGACCCCTG	480
DB	421	GGCATGCGGCGCGGCTGTTCTCATAGTACGTCGCGAAGCGGATCGAGACCCCTG	480
QY	481	ATCGTCCGATGCTGAAGAGCGTCGAGATCCAGTTGGGCACTTCTTCGTGGTCTGACC	540
DB	481	ATCGTCCGATGCTGAAGAGCGTCGAGATCCAGTTGGGCACTTCTTCGTGGTCTGACC	540
QY	541	TACTTCTGATCGTGGCTCGAGATCGGCGAAGCTTACCGAGCGTCTCGAGCGCTG	600
DB	541	TACTTCTGATCGTGGCTCGAGATCGGCGAAGCTTACCGAGCGTCTCGAGCGCTG	600
QY	601	GGCATGATCGCGAGTAAAGTGGTGGCGGCGGCTGGGCACTTCTTCGTACCTGTCGGG	660
DB	601	GGCATGATCGCGAGTAAAGTGGTGGCGGCGGCTGGGCACTTCTTCGTACCTGTCGGG	660
QY	661	AAGTGAAGTTCGCGGAGTACCTGCTGATTCGAACTGACCGGCGCGGCGAGCTGATC	720
DB	661	AAGTGAAGTTCGCGGAGTACCTGCTGATTCGAACTGACCGGCGCGGCGAGCTGATC	720
QY	721	GTGTTCTGCGCGGCTGGTGGCGGCGGCTGGGCACTTCTTCGTACCTGTCGGG	780
DB	721	GTGTTCTGCGCGGCTGGTGGCGGCGGCTGGGCACTTCTTCGTACCTGTCGGG	780
QY	781	GCGCAGTCTTCAATGCGGAGTGGCGGCGGCTGGGCACTTCTTCGTACCTGTCGGG	840
DB	781	GCGCAGTCTTCAATGCGGAGTGGCGGCGGCTGGGCACTTCTTCGTACCTGTCGGG	840
QY	841	GCGTGTATGTCGCCAGGAGTCTGCTGTTTCAATGCGGAGTGGGCTGGGCACTGGA	900
DB	841	GCGTGTATGTCGCCAGGAGTCTGCTGTTTCAATGCGGAGTGGGCTGGGCACTGGA	900

QY	901	ACCTCTCGGTGATGATCCAGTTCGCTTCTTCAAGTGCACGAGCGCGGCTTCCGT	960
DB	901	ACTCTGTGCGTATCCTGCGAGGTGGTCTTCTTAAACTACGCGACAGCGTATTTCCGT	960
QY	961	ATGCGCGGATCCATCACCACTTTTCAAGTTCGAAAGGCTGGCGGACCGCGGTGATCGT	1020
DB	961	ATGCGCGGATCCATCACCACTTTTCAAGTTCGAAAGGCTGGCGGACCGCGGTGATCGT	1020
QY	1021	CGTCTCGGATCATCACCGTATCCTGCTGATCGGCTCGGCTCGGCTGAAAGTTCGT	1080
DB	1021	CGTCTCGGATCATCACCGTATCCTGCTGATCGGCTCGGCTCGGCTGAAAGTTCGT	1080
QY	1081	TGA 1083	
DB	1081	TAA 1083	

RESULT 10  
US-09-815-242-9977  
; Sequence 9977, Application US/09815242  
; Patent No. US20020061569A1  
; GENERAL INFORMATION:  
; APPLICANT: Hasselbeck, Robert  
; APPLICANT: Ohlsen, Kari L.  
; APPLICANT: Zyskind, Judith W.  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John D.  
; APPLICANT: Carr, Grant J.  
; APPLICANT: Yamamoto, Robert T.  
; APPLICANT: Xu, H. Howard  
; TITLE OF INVENTION: Identification of Essential Genes in  
; FILE OF INVENTION: Prokaryotes  
; FILE REFERENCE: ELITRA.011A  
; CURRENT APPLICATION NUMBER: US/09/815,242  
; CURRENT FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: 60/191,078  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: 60/206,848  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR APPLICATION NUMBER: 60/207,727  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: 60/242,578  
; PRIOR FILING DATE: 2000-10-23  
; PRIOR APPLICATION NUMBER: 60/253,625  
; PRIOR FILING DATE: 2000-11-27  
; PRIOR APPLICATION NUMBER: 60/257,931  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: 60/269,308  
; PRIOR FILING DATE: 2001-02-16  
; NUMBER OF SEQ ID NOS: 14110  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 9977  
; LENGTH: 1083  
; TYPE: DNA  
; ORGANISM: Salmonella typhi  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)...(1083)  
; NAME/KEY: misc.feature  
; LOCATION: (1)...(1083)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-815-242-9977

Query Match 47.8%; Score 517.2; DB 9; Length 1083;  
Best Local Similarity 67.3%; Pred. No. 1.1e-124;  
Matches 729; Conservative 0; Mismatches 354; Indels 0; Gaps 0;

QY	1	ATGCTCTGCTGCTGCGGCAATACCTGCAACAGTTCCTACAAGGGCTTCGGGCTCTTCCAG	60
DB	1	ATGTTAGTTGGCTGCGCGAGCAATTTGTCGCAATATATTCCGGCTTTAAACGCTCTTTCT	60
QY	61	TACTGACCTGCGCGGCAATCTTCAAGGCTTCACCGGCTTCGCTCTCGTGGGCTG	120

Db	61	TATCTGACGTTTCGGCCCATCGTCACGCTTTGACCGCGCTGTTCACTCTTTATGGATG	120
Qy	121	GGGCCCTGGATGATCGGTACCTTTGCAGATCCCCCAGATCGGCAGCGCTGCGCAACGAC	180
Db	121	GGCCCGGTATGATCGCTCGTCTGCAAAATCTCTTTTGGCCAGGTGTACGTAAACGAT	180
Qy	181	GGTCCGACGTCGACCTGTGCGAAGAGGCAACCCGACCATGGGCGGCGCCTGATCCCTT	240
Db	181	GGCCCGGATCGGACATTCAGTAAACGGGTACGCCACGATGGCGGCATCATGATCCCTG	240
Qy	241	ACGCCCATAGCCATCAGCACGCTGTCGGCGGATCTTTTCCAAACCGTACGTGTGGTAA	300
Db	241	ACGGGATTTGTATTTCCGTTCTGTTATGGGCTTACCCGTTAAACCCGTACGTCGGTGC	300
Qy	301	GTGCTGTCGTACCTTACCCTGCTGTTGCGTGCCATCGGCTGGGTAGACGACTACCGCAAGTG	360
Db	301	GTGCTGTCGTATTGATCGGCTAGCGCATTAACGTTTGTGATGACATACCGCAAGTG	360
Qy	361	ATCGAGAAGAACTCCCGTGGCTGCCAGCCGTGGAAGTACTTTCTGGCAGTCGGTGTTC	420
Db	361	GTGCGAAAGATACCAAAGGCTGATTCGCCGCTGGAATAATTTCTGAGATGTCGGTTATC	420
Qy	421	GGCATCGGCGCGCGCTGTTCCCTCATGACATGCCGAACCCGATCGAGACCAACCTTG	480
Db	421	CGGCTCGCGCTGGCTTTGGCTTTATCTCGTCGGGAAGACACGCCCGCGACCCACATG	480
Qy	481	ATCGTCCGATGCTGAAGCGTCGAGATCCAGTTGGGCACTTTCTTCTGTGTCGTGACC	540
Db	481	GTGTGCGCTTCTTTAAAGATGTTATGCCCAATTTGGGCTGTTTACATCTCTGCTGCC	540
Qy	541	TACTTCGTCTATCGTCGCTCGAGCAATGCGGTGAACTCATCCGACGCTTCGACGCGCTG	600
Db	541	TACTTTGTCTATCGTCTGCTACGGGTAAACCGCTTAACTTGACCGACGGGCTTGATGTCG	600
Qy	601	CGCATCATCGCAGCGTAAATGGTTGCGCGCGCTGGGCACTTTCTGCTACTGTCGGGC	660
Db	601	CGCATTCGCGACTGCTCTGTTGCGCGCGGCTTTGCGCTGCTGGGCTTGGGCGACCGG	660
Qy	661	AACGTGAAGTTCCGCGAGTACCTGCTCATTTCCAAACGTACCGGCGCGCGAGCTGATC	720
Db	661	AACATGAATTCGCGCAATTAACCTGCATATTCGGTATTTACGCATGCGGCGGAGCTGGT	720
Qy	721	GTGTTCTGCGCGCTGCTGGCGCGCGCTCGGCTTCTCTGTGTTCAACCTATCCG	780
Db	721	ATTGCTGTACGGCGATTGTTCGGCGCGGATTAGGATCTTGTGTTTAAACGCTATCCG	780
Qy	781	CGCAGGCTTTCAATGGCGAGCTCGGCGCGCTGCGCTGGCGCGCGCTGGGACCATC	840
Db	781	CGCAGGTTTTTAATGGCGAGTCCGATCGCTGCGCTTGGCGCGGCTTGGGCAATATC	840
Qy	841	CGGTGTATCTGCGCGAGGAGATCGTGCCTGTTCAATGCGGTGGGTTGCTCATGGAA	900
Db	841	GCCGTGCTGCTCGCTCAGGAATTTCTGCTGTGATCATGGCGCGCTTTTGTGGTGGAA	900
Qy	901	ACCTCTCGGTGATGATCCAGTTCGCTTCCCTCAAGCTGACCGGACCCCGCTTCTCCGT	960
Db	901	ACTCTGTCGGTCACTCTGCAAGTGGTTCTTTTAACTACGCGAGACAGCTATTTCCGT	960
Qy	961	ATGGCGCGATCCATCACATTTGAACTGAAAGGCTGGCGGACCCCGCGCTGATCGTG	1020
Db	961	ATGGCGCTATCCATCACACTATGAACTGAAAGGCTGGCGGAAACCGCGCTGATTTGTG	1020
Qy	1021	CGCTTCTGGATCATCACCGTGATCTCGTGTGATCGGCTCGCCACTTTGAGAGTCCGT	1080
Db	1021	CGCTTCTGGATTTTTCCTGATGCTGTGCTGATTTGGCTGGCAACGCTGAAGTACGT	1080
Qy	1081	TGA	1083
Db	1081	TAA	1083

RESULT 11

RESUB II  
US-09-741-669-258

```

; Sequence 258, Application US/09741669
; Patent No. US20020022718A1
; GENERAL INFORMATION:
; APPLICANT: Forsyth, R. Allyn
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; TITLE OF INVENTION: Genes identified as required for
; TITLE OF INVENTION: proliferation of E. coli
; FILE REFERENCE: ELITRA 009A
; CURRENT APPLICATION NUMBER: US/09/741,669
; CURRENT FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: US 60/173005
; PRIOR FILING DATE: 1999-12-23
; NUMBER OF SEQ ID NOS: 481
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 258
; LENGTH: 1083
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1083)
; US-09-741-669-258

```

Query Match	46.4%	Score 502.2	DB 9	Length 1083
Best Local Similarity	66.5%	Pred. No. 8.6e-121		
Matches 720	Conservative 0	Mismatches 363	Indels 0	Gaps 0
QY	1	ATGCTCTGCTGTGGCCGAATACCTGTCAACAGTCTTACAAGGCGCTTCGCGCTCTCCAG	60	
Db	1	ATGTAGCTTTGGCTGGCCGAACATTGGTCAATATATTTCGGCTTTAACGCTCTTTTC	60	
QY	61	TACCTGACCCCTGGCGGCATCTTCAGSGTCTCACCGCGCTGTGCTGTGCTGTGGCTG	120	
Db	61	TATCTGACGTTTGGCGCATCGTCAGCCTCTCACCGCGCTGTTCATCTCATTTGTGGATG	120	
QY	121	GGGCGCTGGATGATCCGTACTCTTGACATCCCCAGATCGCCAGGCGCTGCGCAACGAC	180	
Db	121	GGCCCGGTATGATGTCTCATTTGCAAAACTTTCCTTTGGTCAGTGGTGGTAACGAC	180	
QY	181	GGTCGCACTGCGCACTGTGCAAGAGGGCAACCCGACCATAGGCGCGCCCTGATCCTT	240	
Db	181	GGTCCTGAATCACACTTTCAGCAAGCGCGGTACGCGCACCATAGGCGCGGATATGATCCTG	240	
QY	241	ACCGCCATAGCCATCAGCAGCGTCTCTGGCGCGGATCTTTCCAAACCGCTACGTGGGTA	300	
Db	241	ACGGCGATTGATCTCCGTACTCTCGGTGGCTTACCCGTCCAATCCGTACGTCGTGTGC	300	
QY	301	GTGCTGTGCTTAACCTGTGTTTCGGTGGCCATCGGCTGGGTAGACGACTACCGAAGGTG	360	
Db	301	GTGTTGTGTGTGCTGTGAGTTACGGTGTATTGGCTTTGTGATGATTATCGCAAGTG	360	
QY	361	ATCGAGAAGAACTCCGTGGCTCCGAGCGCTGGAAAGTACTTCTGGCAGTCGGTGTTC	420	
Db	361	GTGGTAAAGACACAAGGGTTGATCGCTGTTGGAAAGTATTCTGAGTCTCGTCAAT	420	
QY	421	GGCATGCGCGCGCGGTGTTCCCTCTCATATGACTCCGAAACCCCGATCGAGACCCCTG	480	
Db	421	CGCGTGGGTGTGCGCTTCGCGCTGTACGTTCCGCGCAAGACACGCCCGCAACGAGCTG	480	
QY	481	ATCGTCCGATGCTGAGAGCGTCGAGATCCAGTTGGCATCTTCTCTGGTCTCTGACC	540	
Db	481	GTGTTCCATTTTAAAGATGATGATGCCGAGCTGTGGGGTGTTCTACATTTCTGTGGCT	540	
QY	541	TACTTGTGTCATGTGCGGTGAGCAATGCGGTGAACCTCAACGACGGTCTCTGAGCGGCTG	600	
Db	541	TACTTCGTCAITGTGGTACTTGGCAAAGCGGTAAACCTGACCGATGGTCTCGACGGGCTG	600	
QY	601	GGCATCATGCCGACGGTAAATGGTTGCGCGCGCGCTGGGCACTCTTCTGTACTCTGTGGGC	660	
Db	601	GCATTTATGCCGACGATATTGTTGCCCGGTGTTTGGCTGGTGGCTGGGCGACCGCG	660	
QY	661	AACGTGAAGTTGCGCGAGTACCTGCTGATTCCCAACGATACCGGCGCGCGGAGCTGATC	720	



Qy 1021 CGCTTCTGGATCATCACCGTGATCCTGGTGTGATCGGCTCGCCACCTTGAAGCTGGCT 1080  
 Db 1021 CGTTCTGGATTAATTTTCGCTGATGCTGGTCTGATTTGCTGCGAAGCGTGAAGTACGT 1080  
 Qy 1081 TGA 1083  
 Db 1081 TAA 1083

RESULT 13

US-10-282-122A-20228  
 ; Sequence 20228, Application US/10282122A  
 ; Publication No. US20040029129A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Wang, Liangsu  
 ; APPLICANT: Zamudio, Carlos  
 ; APPLICANT: Malone, Cheryl  
 ; APPLICANT: Haselbeck, Robert  
 ; APPLICANT: Ohlsen, Kari  
 ; APPLICANT: Zyskind, Judith  
 ; APPLICANT: Wall, Daniel  
 ; APPLICANT: Trawick, John  
 ; APPLICANT: Carr, Grant  
 ; APPLICANT: Yamamoto, Robert  
 ; APPLICANT: Forsyth, R.  
 ; APPLICANT: Xu, H.  
 ; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
 ; FILE REFERENCE: ELITRA.034A  
 ; CURRENT APPLICATION NUMBER: US/10/282.122A  
 ; CURRENT FILING DATE: 2003-02-20  
 ; PRIOR APPLICATION NUMBER: 60/191,078  
 ; PRIOR FILING DATE: 2000-03-21  
 ; PRIOR APPLICATION NUMBER: 60/206,848  
 ; PRIOR FILING DATE: 2000-05-23  
 ; PRIOR APPLICATION NUMBER: 60/207,727  
 ; PRIOR FILING DATE: 2000-05-26  
 ; PRIOR APPLICATION NUMBER: 60/230,335  
 ; PRIOR FILING DATE: 2000-09-06  
 ; PRIOR APPLICATION NUMBER: 60/230,347  
 ; PRIOR FILING DATE: 2000-09-09  
 ; PRIOR APPLICATION NUMBER: 60/242,578  
 ; PRIOR FILING DATE: 2000-10-23  
 ; PRIOR APPLICATION NUMBER: 60/253,625  
 ; PRIOR FILING DATE: 2000-11-27  
 ; PRIOR APPLICATION NUMBER: 60/257,931  
 ; PRIOR FILING DATE: 2000-12-22  
 ; PRIOR APPLICATION NUMBER: 60/267,636  
 ; PRIOR FILING DATE: 2001-02-09  
 ; PRIOR APPLICATION NUMBER: 60/269,308  
 ; PRIOR FILING DATE: 2001-02-16  
 ; Remaining Prior Application data removed - See File Wrapper or PALM.  
 ; NUMBER OF SEQ ID NOS: 78614  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 20228  
 ; LENGTH: 1083  
 ; TYPE: DNA  
 ; ORGANISM: Escherichia coli  
 ; US-10-282-122A-20228

Query Match 45.4%; Score 502.2; DB 13; Length 1083;  
 Best Local Similarity 66.5%; Pred. No. 8.6e-121; Indels 0; Gaps 0;  
 Matches 720; Conservative 0; Mismatches 363;  
 Qy 1 ATGCTCTGCTGCTGGCGGAATACCTGCAACAGTTCTACAGGGCTTCGGCGTCTTCCAG 60  
 Db 1 ATGTTAGTTTGGCTGGCGGAACATTTGGTCAAAATATTATTCCGGCTTTAAACGTTTTC 60  
 Qy 61 TACTGTACCTTGGCGGATTTCTCAGGTGCTACCGGCTGCTGCTGCTGCTGCTGCTG 120  
 Db 61 TATCTGAGTTTGGCGGATTCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120  
 Qy 121 GGGCCCTGGATGATCCGTACCTTGCAGATCCCCAGATCGGCCAGCGGCTGCGCAACG 180

RESULT 14

US-10-282-122A-37049  
 ; Sequence 37049, Application US/10282122A  
 ; Publication No. US20040029129A1  
 ; GENERAL INFORMATION:

Db 121 GGGCCCGGTATGATTTGCTCATTTTGCACAAAATCTTTCTTTGGTGGTGGTGGTGGTGGT 180  
 Qy 181 GGTCCGAGTCCGACCTGTGCAAGAGGGCAACCCGACCATGGGCGGCGGCGGCGGCGGCGG 240  
 Db 181 GGTCTGTAATCACACTTCAGCAAGCGGCTACGCGGACCATGGGCGGGAATATGATCTG 240  
 Qy 241 ACCGCCATAGCCATCAGCACGCTCTGTGGCGGATCTTTCCAAACCGCTAGCTGGGTA 300  
 Db 241 ACCGCGATTGTGATCTCGTACTGCTGGGCTTACCGCTCCAATCGTAGCTCTGGTGC 300  
 Qy 301 GTGCTGGTCCGTACCCCTGCTTGGTGCATCGGCTGGGTAGAGTACGACTACCGAAGTG 360  
 Db 301 GTGTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 360  
 Qy 361 ATCGAGAAAGAACTCCCGTGGCTCGCGAGCGCTGGAAGTACTTTCTGGCAGCTCGGTTC 420  
 Db 361 GTGCGTAAAGACACCAAGGGTGTGATCGCTGTTGGAAGTATTTCTGGATCGGTCAAT 420  
 Qy 421 GGCATCGGCGGCGGCTGTTCTCTATGATCTCCGAAACCCCGATCGAGACCACTG 480  
 Db 421 GCGCTGGTGTGGCTTCGCCCTGTACTTCCGCGCAAGACACGCCCGCAACGAGCTG 480  
 Qy 481 ATCGTCCGATGCTGAAGAGCGTCCAGATCCAGTTGGGCATCTTTCTTGGTGGTCTGACC 540  
 Db 481 GTGGTCCCATCTTTAAAGATGTGATCGCGCAGCTGGGCTGTTCTACATTTCTGGCT 540  
 Qy 541 TACTTCTGATCGTCCGCTCGAGCAATCGGCTGAACTCCACCGAGCTCTCGAGGCTG 600  
 Db 541 TACTTCTGATTTGGGCTTCTGGCAACCGGTAACCTGACCGGATGGTCTCGAGGCTG 600  
 Qy 601 GCGATCATGCCGAGCGGTAAATGGTTGCCGCGCTGGGCACTTTCTGTACTCTGCGGC 660  
 Db 601 GCAATATGCCGACCGTATTTGTGCGCGTGGTTTGGCTGGTGGGCTGGGCGACCGCG 660  
 Qy 661 AACGTGAAGTTGCCGAGTACTGCTGATTCGACGATCCGCGGCGGCGGCGGCGGAGCTG 720  
 Db 661 AATATGAACCTTTGCCAGCTACTTGCATATACCGTATCTGCGACACGCGCGGGAAC 720  
 Qy 721 GTGTTCTGCGCGGCTGCTGCGCGCGCTCGGCTTCTCTGTTCAACACCTATCCG 780  
 Db 721 ATTGTCTGATCCGCGATAGTGGGCGAGGATGGGCTTCTGTTTAAACACCTATCCG 780  
 Qy 781 GCGCAGGTCTTCAATGGGCGAGTGGGCGCTGGGCGCTGGGCGGCGGCGGCGGCGGCGG 840  
 Db 781 GCGCAGGTCTTATGGGCGATGATAGTTTCGCTGGCTTAGGTGGTGGCTTAGGCATTA 840  
 Qy 841 GCGGTGATCGTGGCGCAGGAGATCGTCTGTTTCATGCTGGTGGGCTGTTCTGTCATGAA 900  
 Db 841 GCGGTACTGCTACGTCAGGAATTCCTGCTGGTATTATGGGCGGCTGTTCTGTTAGTA 900  
 Qy 901 ACCCTCTCGGTGATGATCCAGTCCGCTTCTTCAAGCTGACCGGAGCGGCGGCTCTCGT 960  
 Db 901 ACGCTTCTGTCATCTCTGAGTCCGCTTAAAGCTGGCGGACACGATATTTTCGCG 960  
 Qy 961 ATGGCGCGATCCATCCATCCATTCGAACTGAAAGCTGGCGGACCCGCGGCTGATCCG 1020  
 Db 961 ATGGCACCGATTTCATCCATCCATGATGAAAGCTGGCGGAAACCCGCGGCTGATTTG 1020  
 Qy 1021 CGCTTCTGGATCATCACCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080  
 Db 1021 CGTTCTGGATTTTCGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080  
 Qy 1081 TGA 1083  
 Db 1081 TAA 1083

```

; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITEA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 37049
; LENGTH: 1060
; TYPE: DNA
; ORGANISM: Salmonella paratyphi A
US-10-282-122A-37049

Query Match 45.3%; Score 490.6; DB 13; Length 1060;
Best Local Similarity 68.2%; Pred. No. 8.9e-118;
Matches 710; Conservative 0; Mismatches 329; Indels 2; Gaps 2;

QY 43 GGTTCGGGCTTTCAGTACCTGCGGCGCATCTAGCGTGTCTACCGCGCTG 102
DB 19 GGTTCGAGCTCTTTCTTATCTGACGTTTCGGCGCATCTGAGCCTGTGACCGGCTG 78
QY 103 TCCTGTCTGCTGTGGCTGGGGCCCTCGATGATCCGTACCTTGAGATCCCCAGATCGGC 162
DB 79 TTCACTCTTTATGATGGCCCGGTATGATCGCTCGTGCAGAAATCTCTTTTGGC 138
QY 163 CAGCGCTGGCGAAGCAGCGTTCGCGAGTCGACCTGTGAGAGAGGCGACCCGACCATG 222
DB 139 CAGGTGTGTACGTAACTGATGCGCGGATCGCACTTCAGTAAACGCGGTACGCGCATG 198
QY 223 GCGCGGCGCTGTCTTACCGCATAGCATCAGACGCTGCTGGGCGGATCTTTCC 282
DB 199 GCGGATCATGATCTGACGGCGATGATGATTCGTTCTGTATGGGCTTACCGCT 258
QY 283 AACCGGTACGTGGGTAGTGTGCTGTGTTACCTGCTGCTGCGTCCCATCGGCTGGGTA 342
DB 259 AACCGGTACGTGGGTGCTGTGTTGTTATGATTGGGTACGGCATTTATCGTTTTCG 318
QY 343 GAGCACTACCGAAGTGTACGAGAGACTCCCGTGGCTCGGAGCGCTGGAGTAC 402
DB 319 GATGACTACCACAAAGTGTGGGAAAGATACCAAAAGGCTGATTGCGGCTGGAAATAT 378

403 TTCTGGCAGTCGGTGTTCGGCATCGGCGCGCGTGTCTCTACATGACTGCGGAACC 462
379 TTCTGGATGTCGGTTATCGCGCTCGG-GTGGCCCTTTTCGGCTTTATCTCTCGGGAAGAC 437
463 CCGATCGAGACCACTGATCGTGGCGCATGCTGAAGAGCGTTCGAGATCCAGTTGGGATC 522
438 AGCCCGCGACCAACTGGTGGTGGTCTTTTAAAGATGTTATGCCGCAATTTGGGCTG 497
523 TTCTCGTGTCTGACCTACTTCCTGATCGTGGGTTCGAGCAATGCGGTGAACCTCAAC 582
498 TTTTACATTTCTGCTCTACTTTTGTTCATCGTCGATACCGGTAAACCGCTGAAAC 557
583 GAGGCTCTCACGCGCTGGCGATCATGCCGAGCGTAAATGTTGCGCGCGCTGGGATC 642
558 GACGGGCTTGATGGTCTGGCGATATGCCGACTGTTTTTCGTTGCGCGCGCTGGGCTG 617
643 TTCTGTACTCTGTGGGCAAGTGAAGTTCGCGAGTACCTGTGATTCACACGATCCG 702
618 GTGGCTTGGGCGACCGGGAACATGAACTTCGCAATTTACCTGCATATTCGCTATTACGC 677
703 GCGCGCGCGAGTGTGCTGCGCGCGCTGTCGCGCGCGCTGTCGCGCGCGCTGCGCTC 762
678 TATCGGGCGAGCTGGTGTGTTGCTGACGGCGATTTGTCGCGCGCGGATTTAGATTCTG 737
763 TGGTTCAACACCTATTCGCGCGCGAGTCTTCATGGGCGACGTCGCGCGCTGGGCTGGGC 822
738 TGGTTTAAACACTATTCGCGCGCGAGTCTTTTATGGGCGACGTCGATCGCTGGGCTGGGC 797
823 GCGCGCTGGGCGACCATCGGCTGATCGTGGCGCGAGATCGTGGCGCGAGATCGTGTTCATCATGGCT 882
798 GCGCGCTGGGCGCATTTATCGCGCGCTGCTGCTGCGCAGAGTTCCTGCTGGTGTGATCATGGC 857
883 GGGGTGTTCTGTCATGGAACCCCTCTCGTGATGATCCAGGTGCTTCTTCAAGCTGACC 942
858 GGGCTTTTGTGGTGGAACTCTGTCGTCATCTCGCAGGTGGTTCCTTTAAACTACGC 917
943 GGAGCGCGCTTTCGCTATGGCGCGCATCATCATCACCATTTCGAACTGAAAGCTGGCG 1002
918 GGACAGC-TATTTTCCGTATGGCGCTTATCCATCACCATTGAACTGAAAGCTGGCG 976
1003 GACCGCGCGTGTGTCGCTGTCGATCATCACCGCTGATCCCTGCTGCTGATCGGCTC 1062
977 GAACCGCGGTGATTTGCGCTTCTGATTTATTCGCTGATGCTGCTGATGGCTG 1036
1063 GCCACCTTGAAGTGGCTTGA 1083
1037 GCAACGCTGAAGGTACGTAA 1057

RESULT 15
US-10-282-122A-15193
; Sequence 15193, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITEA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
```

; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 15193
; LENGTH: 1167
; TYPE: DNA
; ORGANISM: Bordetella pertussis
US-10-282-122A-15193

Query Match 44.4%; Score 480.4; DB 13; Length 1167;
Best Local Similarity 66.7%; Pred. No. 4e-115;
Matches 778; Conservative 0; Mismatches 301; Indels 87; Gaps 3;

QY 1 ATGCTCTGCTGCGCGGATCTGCAAGTCTTACAGGGCTTCGGGCTTCCAG 60
DB 1 ATGCTGCTTGAGATCGCCCGGTGGCTGCCAGACGATGCGCGCATTTGGGATTCGAG 60
QY 61 TACCTGAGCCCTGCGCGGATCTCAGCGTGTCTACCGCGCTGTGCTGTGCTGTGCTG 120
DB 61 TACATCAGCTGCGCGCGCTGTGCGTGTGCGGCGCTGCTGATCGGCTGTGCGG 120
QY 121 GGGCCCTGATGATCGGTACCTTGCAGATCCCCAGATCGCCAGCGCTGCGCAAGCAG 180
DB 121 GGGCGCGCGTGTATCGCGCGCTGACCGAAATGAAGATCGGCCAGCGCGCTGCGCGCTAC 180
QY 181 GGTCCGCGAGTCCGACCTGTGCAAGAGGCGACCCGACCATGGGCGCGCGCTGATCCTT 240
DB 181 GGCCTGGAGTCCGACCTGTGCAAGAGGCGACCCGACCATGGGCGCGCGCTGATCCTT 240
QY 241 ACCGCGATAGCCATCAGCAGCTGTGTGGGCGGATCTTTCAACCGCTACGTGTGGTA 300
DB 241 ATCGCGATCGCCATCAGCAGCTGTGTGGGCGGATCTTTCAACCGCTACGTGTGGTA 300
QY 301 GTGCTGTGCTGAGCTTGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360
DB 301 GTGCTGTGCTGAGCTTGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360
QY 361 ATCGAGAGAACTCCCGTGGCTGCGGCGCTGCGGCGCTGCGGCGCTGCGGCGCTGCGG 420
DB 361 GTCTACCGCGATCCCGAGGCGATCGCGCGCTGCGGCGCTGCGGCGCTGCGGCGCTGCGG 420
QY 421 GGCATCGGCGCGCGCTGCTCTACATGATGCTGCGGAAAC----- 462
DB 421 GGGCTGTGCGCGCGCTGCTCTACATGATGCTGCGGAAAC----- 462
QY 463 -----CCGATCGAGACC 474
DB 463 -----CCGATCGAGACC 474
QY 481 TGGCGCTCTTCAAGGATGGGTGGGTAGCGGTTTCAAGATCGCGCTGCGGCGCGGCGCC 540
DB 481 TGGCGCTCTTCAAGGATGGGTGGGTAGCGGTTTCAAGATCGCGCTGCGGCGCGGCGCC 540
QY 475 ACCCTGATCGTCCGATGCTGAGAGCTGCGAGATCGAGTGGGATCTT---CTTCGTG 531
DB 541 GACCTGATCGTCCGATGCTGAGAGCTGCGAGATCGAGTGGGATCTT---CTTCGTG 600
QY 532 GTCTGACCTACTTCTGCTGCTGCGCTGCGGCGCTGCGGCGCTGCGGCGCTGCGGCGCTG 591
DB 601 GGCCTGACCTGCGGCGCTGCTGCTGCGGCGCTGCGGCGCTGCGGCGCTGCGGCGCTG 660

Search completed: May 9, 2004, 15:56:26
Job time : 531 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 9, 2004, 13:20:48 ; Search time 97 Seconds  
(without alignments)  
6195.998 Million cell updates/sec

Title: US-10-089-787-1  
Perfect score: 1083  
Sequence: 1 atgctccgtgctggtgcca.....ccacctgaagctggtga 1083  
Scoring table: IDENTITY NUC  
Gapop 10\_0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents NA:\*  
1: /cgn2\_6/ptodata/2/ina/5A\_COMB.seq:\*  
2: /cgn2\_6/ptodata/2/ina/5B\_COMB.seq:\*  
3: /cgn2\_6/ptodata/2/ina/6A\_COMB.seq:\*  
4: /cgn2\_6/ptodata/2/ina/6B\_COMB.seq:\*  
5: /cgn2\_6/ptodata/2/ina/PTUS\_COMB.seq:\*  
6: /cgn2\_6/ptodata/2/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	1068.6	98.7	1170	4	US-09-252-991A-7929
2	1068.6	98.7	1404	4	US-09-252-991A-7623
3	455	42.0	1032	4	US-09-489-039A-5268
4	366.2	33.8	1083	4	US-09-543-681A-2879
5	361.4	33.4	1830121	4	US-09-557-884-1
6	361.4	33.4	1830121	4	US-09-643-990A-1
7	250.6	23.1	1119	4	US-09-328-352-2043
8	233.8	21.6	1107	4	US-09-340-236-1218
9	232.2	21.4	100848	4	US-09-596-002-39
C 10	175.4	16.2	4403765	3	US-09-103-840A-2
C 11	175.4	16.2	4411529	3	US-09-103-840A-1
C 12	174.2	16.1	640681	4	US-09-790-988-1
13	134	12.4	210	4	US-09-489-039A-5281
14	133.4	12.3	6911	1	US-08-311-174-4
15	115.2	10.6	2402	4	US-09-221-017B-881
16	91.2	8.4	1230025	4	US-09-198-452A-1
17	88.4	8.2	996	4	US-09-107-532A-1622
18	87.4	8.1	981	3	US-08-986-768-1
19	87.4	8.1	981	3	US-08-986-768-3
20	87.4	8.1	3190	3	US-08-986-768-4
21	86	7.9	8148	4	US-08-961-527-11
C 22	79.4	7.3	1600	4	US-08-956-171B-159
23	79.4	7.3	2423	3	US-08-714-918-86
24	79.4	7.3	2423	3	US-09-265-315-86
25	79.4	7.3	2423	3	US-09-265-315-86
26	79.4	7.3	2423	3	US-09-266-417-86
27	79.4	7.3	2423	4	US-09-528-709-86

28	79.4	7.3	2423	4	US-09-527-745-86	Sequence 86, Appl
29	73.2	6.8	1005	4	US-09-134-001C-1014	Sequence 1014, Ap
30	72.2	6.7	984	4	US-09-134-000C-1280	Sequence 1280, Ap
31	71.4	6.6	1416	4	US-09-489-039A-5247	Sequence 5247, Ap
32	70.4	6.5	330	3	US-09-197-649-7	Sequence 7, Appli
C 33	69.2	6.4	1926	4	US-09-249-585A-4	Sequence 4, Appli
34	69.2	6.4	1931	2	US-09-130-114-2	Sequence 2, Appli
35	68.6	6.3	1338	4	US-09-252-991A-16571	Sequence 16571, A
C 36	68.6	6.3	1404	4	US-09-252-991A-16031	Sequence 16031, A
37	68.6	6.3	2157	4	US-09-252-991A-16462	Sequence 16462, A
38	68	6.3	4403765	3	US-09-103-840A-2	Sequence 2, Appli
39	68	6.3	4411529	3	US-09-103-840A-1	Sequence 1, Appli
40	64.4	5.9	1392	4	US-09-252-991A-8917	Sequence 8917, Ap
41	64.4	5.9	1617	4	US-09-252-991A-8802	Sequence 8802, Ap
C 42	64.4	5.9	1956	4	US-09-252-991A-9165	Sequence 9165, Ap
43	63	5.8	1995	4	US-09-252-991A-14545	Sequence 14545, A
44	63	5.8	3396	4	US-09-252-991A-14676	Sequence 14676, A
C 45	63	5.8	3444	4	US-09-252-991A-15078	Sequence 15078, A

ALIGNMENTS

RESULT 1  
US-09-252-991A-7929/c  
; Sequence 7929, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196-136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 7929  
; LENGTH: 1170  
; TYPE: DNA  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-7929

Query Match	98.7%;	Score	1068.6;	DB 4;	Length	1170;
Best Local Similarity	99.2%;	Pred. No.	2.1e-23;			
Matches 1074;	Conservative	0;	Mismatches	9;	Indels	0;
Gaps	0;					
1	ATGCTCTGCTGCTGGCGGAATACCTGCAACAGTTCTACAAAGGGCTTCGGCGTCTTCCAG	60				
1167	ATGCTCTGCTGCTGGCGGAATACCTGCAACAGTTCTACAAAGGGCTTCGGCGTCTTCCAG	1108				
61	TACCTGACCTTGC CGGCATCTCAGCGTGCTCACCGCGCTGTGCGCTGTGCGTGTGGGTG	120				
1107	TACCTGACCTTGC CGGCATCTCAGCGTGCTCACCGCGCTGTGCGCTGTGCGTGTGGGTG	1048				
121	GGGCGCTGGATGCTGATCTTCAGATCCCGGAGGCGGCGTGGCAACGAC	180				
1047	GGGCGCTGGATGCTGATCTTCAGATCCCGGAGGCGGCGTGGCAACGAC	988				
181	GCTCGGAGTGCACCTGTGCAAGAGGCGGCGGCGGCGGCTGATCTT	240				
987	GCTCGGAGTGCACCTGTGCAAGAGGCGGCGGCGGCGGCTGATCTT	928				
241	ACCGGCATAGCCATCAGCACGCTGTGTGGCGGATCTTTCGAACCGCTACGTGTGGGTA	300				
927	ACCGGCATAGCCATCAGCACGCTGTGTGGCGGATCTTTCGAACCGCTACGTGTGGGTA	868				
301	GTGCTGTGCTGTTACCTGCTGCTTGGTGCCATCGGCTGGGTAGACGACTACCGCAAGGTG	360				
867	GTGCTGTGCTGTTACCTGCTGCTTGGTGCCATCGGCTGGGTAGACGACTACCGCAAGGTG	808				

361 ATCGAGAGAACTCCCGTGGCTCCGAGCGCTGGAGTACTTCTGGCAGTGGGTTC 420  
Db ATCGAGAGAACTCCCGTGGCTCCGAGCGCTGGAGTACTTCTGGCAGTGGGTTC 748  
Qy 421 GGCATCGCGCGCGCGGTTCCTCTACATGACTCCGAAACCCGATCGAGACACCCCTG 480  
Db 747 GGCATCGCGCGCGCGGTTCCTCTACATGACTCCGAAACCCGATCGAGACACCCCTG 688  
Qy 481 ATCTGTCGGATGCTGAAGAGGTGAGATCCAGTTGGGCACTTCTTCTGGTGGTCTGAC 540  
Db 687 ATCTGTCGGATGCTGAAGAGGTGAGATCCAGTTGGGCACTTCTTCTGGTGGTCTGAC 628  
Qy 541 TACTTCTGTCATCGTGGCTCGAGCAATCGGTAACCTTCCGACGCTCTCGACGCGCTG 600  
Db 627 TACTTCTGTCATCGTGGCTCGAGCAATCGAGTAACTTCCGACGCTCTCGACGCGCTG 568  
Qy 601 GCGATCATGCGAGCGGTAACTGGTTCGCGCGCGCTGGGCACTTCTCTGCTACCTGTGGGC 660  
Db 567 GCGATCATGCGAGCGGTAACTGGTTCGCGCGCGCTGGGCACTTCTCTGCTACCTGTGGGC 508  
Qy 661 AACGTGAAGTTCCGCGAGTACTGCTGATTTCCCAACGTACCGGCGCGCGCGAGCTGATC 720  
Db 507 AACGTGAAGTTCCGCGAGTACTGCTGATTTCCCAACGTACCGGCGCGCGCGAGCTGATC 448  
Qy 721 GTGTTCTGCGCGCGCTGGTCCGCGCGCGCTGGGCACTTCTCTGTTTCAACCTATCCG 780  
Db 447 GTGTTCTGCGCGCGCTGGTCCGCGCGCGCTGGGCACTTCTCTGTTTCAACCTATCCG 388  
Qy 781 GCGAGGTCTTCAATGGGCGAGCTGGCGCGCTGGGCACTTCTGGGCGCGCGCGAGCTGATC 840  
Db 387 GCGAGGTCTTCAATGGGCGAGCTGGCGCGCTGGGCACTTCTGGGCGCGCGCGAGCTGATC 328  
Qy 841 GCGGTGATGCTGGCGAGGATGCTGCTGTTTCAATGAGTGGGCGCGCGCGAGCTGATGATG 900  
Db 327 GCGGTGATGCTGGCGAGGATGCTGCTGTTTCAATGAGTGGGCGCGCGCGAGCTGATGATG 268  
Qy 901 ACCCTCTCGGTGATGATCCAGTCCCTTCCCTTCAAGCTGACCGGACGCGCGCTTCCGT 960  
Db 267 ACCCTCTCGGTGATGATCCAGTCCCTTCCCTTCAAGCTGACCGGACGCGCGCTTCCGT 208  
Qy 961 ATGGCGCGGATCATACATTTGAACTGAAAGGTGGCGCGCGCGAGCTGATGATGATG 1020  
Db 207 ATGGCGCGGATCATACATTTGAACTGAAAGGTGGCGCGCGCGAGCTGATGATGATG 148  
Qy 1021 GCCTTCTGATCATCACCGTATCTGCTGATCGGCTTCCGACCTTGAAGCTGCTG 1080  
Db 147 GCCTTCTGATCATCACCGTATCTGCTGATCGGCTTCCGACCTTGAAGCTGCTG 88  
Qy 1081 TGA 1083  
Db 87 TGA 85

US-09-252-991A-7623  
Query Match 98.7%; Score 1068.6; DB 4; Length 1404;  
Best Local Similarity 99.2%; Pred. No. 2.2e-232;  
Matches 1074; Conservative 0; Mismatches 9; Indels 0; Gaps 0;  
Qy 1 ATGCTCTCTGCTGCTGGCGGAATACCTGCAACAGTCTTCAAGAGGCTTCGGGCTCTTCAG 60  
Db 322 ATGCTCTCTGCTGCTGGCGGAATACCTGCAACAGTCTTCAAGAGGCTTCGGGCTCTTCAG 381  
Qy 61 TACCTGACCTTGGCGGCACTTCTGAGCGTGTCTACCGGCTGTCTGCTGCTGCTGCTGCTG 120  
Db 382 TACCTGACCTTGGCGGCACTTCTGAGCGTGTCTACCGGCTGTCTGCTGCTGCTGCTGCTG 441  
Qy 121 GGGCCCTGATGATCCGTAACCTTTCAGATCCCGGAGATCGGCGAGCCGCTGGCGAACAC 180  
Db 442 GGGCCCTGATGATCCGTAACCTTTCAGATCCCGGAGATCGGCGAGCCGCTGGCGAACAC 501  
Qy 181 GGTTCGCGAGTGCACCTGTGGAAGAGGCGACCCCGACCATGGGGGGCGCCCTGATCCTT 240  
Db 502 GGTTCGCGAGTGCACCTGTGGAAGAGGCGACCCCGACCATGGGGGGCGCCCTGATCCTT 561  
Qy 241 ACCGCCATAGCCATCAGCACGCTGCTGGGCGGATCTTTCACCCGCTACGTTGGGTA 300  
Db 562 ACCGCCATAGCCATCAGCACGCTGCTGGGCGGATCTTTCACCCGCTACGTTGGGTA 621  
Qy 301 GTGCTGGTGGTACCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360  
Db 622 GTGCTGGTGGTACCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 681  
Qy 361 ATCGAGAGAACTCCCGTGGCTGCGGAGCGCTGCGAGCGCTGCGAGCTGCTGCGAGTCTT 420  
Db 682 ATCGAGAGAACTCCCGTGGCTGCGGAGCGCTGCGAGCGCTGCGAGCTGCTGCGAGTCTT 741  
Qy 421 GGCATCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480  
Db 742 GGCATCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 801  
Qy 481 ATCTGTCGCTGCTGAAAGAGCTCGAGATCCAGTCTGGGCACTTCTTCTGCTGCTGCTG 540  
Db 802 ATCTGTCGCTGCTGAAAGAGCTCGAGATCCAGTCTGGGCACTTCTTCTGCTGCTGCTG 861  
Qy 541 TACTTCTGCTGCTGCGGCTGAGAGCAATCGGCTGAACTCACCAGAGCTTCCAGCGGCTG 600  
Db 862 TACTTCTGCTGCTGCGGCTGAGAGCAATCGGCTGAACTCACCAGAGCTTCCAGCGGCTG 921  
Qy 601 GGCATCATCCGAGCTGATGCTGCTGATTTCCGCGCGCTGGGCACTTCTGCTGCTGCTG 660  
Db 922 GGCATCATCCGAGCTGATGCTGCTGATTTCCGCGCGCTGGGCACTTCTTCTGCTGCTG 981  
Qy 661 AACGTGAAGTTCCGCGAGTACCTGCTGATTTCCGCGCGCTGGGCACTTCTTCTGCTGCTG 720  
Db 982 AACGTGAAGTTCCGCGAGTACCTGCTGATTTCCGCGCGCTGGGCACTTCTTCTGCTGCTG 1041  
Qy 721 GTGTTCTGCGCGCGCTGCTGCGCGCGCTGCGGCTTCTGCTGCTGCTGCTGCTGCTGCTG 780  
Db 1042 GTGTTCTGCGCGCGCTGCTGCGCGCGCTGCGGCTTCTGCTGCTGCTGCTGCTGCTGCTG 1101  
Qy 781 GCGCAGGCTTCTCATGCGGCACTGCGCGCGCTGCGGCTGCGGCGCGCTGCGGCGCGCTG 840  
Db 1102 GCGCAGGCTTCTCATGCGGCACTGCGCGCGCTGCGGCTTCTGCGGCGCGCTGCGGCGCGCT 1161  
Qy 841 GCGGTGATGCTGCGCGCGAGGATCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900  
Db 1162 GCGGTGATGCTGCGCGCGAGGAAATCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1221  
Qy 901 ACCCTCTCGGTGATGATCCAGTTCGCTTCTTCAAGCTGACCGGAGCGCGCGCTTCTCGGT 960  
Db 1222 ACCCTCTCGGTGATGATCCAGTTCGCTTCTTCAAGCTGACCGGAGCGCGCGCTTCTCGGT 1281  
Qy 961 ATGGCGCGGATTCATCACCATTTCGAAGTGAAGGCTGCGGCGCGCGCGGCTGATGCTG 1020  
Db 1282 ATGGCGCGGATTCATCACCATTTCGAAGTGAAGGCTGCGGCGCGCGCGGCTGATGCTG 1341

RESULT 2  
US-09-252-991A-7623  
; Sequence 7623, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; PRIOR FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 7623  
; LENGTH: 1404  
; TYPE: DNA  
; ORGANISM: Pseudomonas aeruginosa

QY 1021 CGCTTCTGGATCATACCGTGTGCTGCTGATCGCGCTCGCCACCTTGAAGCTGGT 1080  
Db |||||  
QY 1342 CGCTTCTGGATCATACCGTGTGCTGCTGATCGCGCTCGCCACCTTGAAGCTGGT 1401  
Db |||||

QY 1081 TGA 1083  
Db 1402 TGA 1404

## RESULT 3

US-09-489-039A-5268  
; Sequence 5268, Application US/09489039A  
; Patent No. 6610836  
; GENERAL INFORMATION:  
; APPLICANT: Gary Breton et. al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA  
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 2709.2004001  
; CURRENT APPLICATION NUMBER: US/09/489,039A  
; PRIOR FILING DATE: 2000-01-27  
; PRIOR APPLICATION NUMBER: US 60/117,747  
; PRIOR FILING DATE: 1999-01-29  
; NUMBER OF SEQ ID NOS: 14342  
; SEQ ID NO 5268  
; LENGTH: 1032  
; TYPE: DNA  
; ORGANISM: Klebsiella pneumoniae  
US-09-489-039A-5268

Query Match 42.0%; Score 455; DB 4; Length 1032;  
Best Local Similarity 67.0%; Pred. No. 6.3e-94;  
Matches 661; Conservative 0; Mismatches 325; Indels 1; Gaps 1;

QY 1 ATGCTCCTGTGTGCGCGAATACCTGCAACAGTTCTACAAGGGCTTCGGCGCTTCCAG 60  
Db |||||  
QY 46 ATGTTAGTATGGTGGCGCAACTTGGTCAMATATTAATTCGGCTTTAAAGCTTTTCT 105  
Db |||||  
QY 61 TACCTGACCTGGCGGCAATCTCAGGTGCTCAGCGCTGTGCGTGTGCTGTGGCTG 120  
Db |||||  
QY 106 TATCTGACGTTTCGCGCCATGFCAGCCCTGTGACCGCTGTGTCATCTCGTTGGATG 165  
Db |||||  
QY 121 GGGCCCTGTGATCGGTACCTTTCAGATCCCGGATCGGCGAGCCGTGGCAACGAC 180  
Db |||||  
QY 166 GGGCCGCGATGATCGCCGCTCTGCAAAACTCGCCCTTTGGCAGGTCTGATAGAC 225  
Db |||||  
QY 181 GGTCCGAGTCGACCTGTGGAAGAGGACCCCGACCATGGCGCGCCCTGATCCTT 240  
Db |||||  
QY 226 GGGCCGAGTCTCATTTTCAGTAAACGCGTACTCCGACCATGGCGGATCATGATCCTC 285  
Db |||||  
QY 241 ACCGCCATAGCCATCAGCAGCTGTGTGGGGATCTTTCCAAACGCTACGTGTGGTA 300  
Db |||||  
QY 286 ACCCGATCACCGTTTCGCTGTGGGCTATCCATCTAACCCGTACGTCTGTGGTC 345  
Db |||||  
QY 301 GTGCTGTGCTTACCTGTGTGCTGTGCGATCGGTGGGTAGACACTACCGCAAGGTG 360  
Db |||||  
QY 346 GTACTGACGGTATTAATCGGCTACGGCATCATCGGTTTGTGATGATTACCGTAAAGTC 405  
Db |||||  
QY 361 ATCGAGAGAACTCCCGTGGCTCCGAGCGCTGGAAGTACTTTCGCAAGTGGGTTC 420  
Db |||||  
QY 406 GTGGGAAAGATACCAAGAGCCCTGATCGCCCGTGGAAATTTTCGATGCGGTGATC 465  
Db |||||  
QY 421 GGATCGGCGCGCGCTGTGTTCTCTACATGACTCGCGAAACCCGATCGAGACCAACCTG 480  
Db |||||  
QY 466 GCGCTGGCGTGGCTTCGGCGTGTATCTGGCGGGGAAAGATACCCGCAACCGAGCTG 525  
Db |||||  
QY 481 ATGCTGCGATGCTGAAGCGTTCGAGATCCAGTTGGGATCTTCTTCGTGTGCTGAC 540  
Db |||||  
QY 526 GTGGTGGCGGTCTTTAAAGACGTAATGCCGACGCTGGGCTGTCTATATCTTGTGGCC 585  
Db |||||  
QY 541 TACTTCTGTCTCGTGGCTCGAGCAATGCGGTGAACCTCAGCAGCGTCTCGACGCGCTG 600  
Db |||||  
QY 586 TACTTCTGTCTCGTGGTACCGGACGCGGTCAACCTGACCGACGCGCTCGACGCGCTG 645  
Db |||||

QY 601 GCGATCATCGGCGAGTAATGCTTGGCGCGCGCTGGGCATCTTCTGCTACCTGTGCGGC 660  
Db |||||  
QY 646 GCGATTATGCGGACCGTTTCGTCGCGGAGGCTTTGCGTGGTGGCTTGGGCCACCGGT 705  
Db |||||  
QY 661 AACGTGAAGTTCGCGAGTACCTGCTGATTTCCAAAGTACCGGGCGCGCGAGCTGATC 720  
Db |||||  
QY 706 AACATGAATTCGCCAACTATCTGCATATCCGTATCTGCGCCACGCGCGAGCTGGT 765  
Db |||||  
QY 721 GTGTTCTGCGCGCTGTGCGCGCGCTGGCTCTGCTGTTCAACACCTATCCG 780  
Db |||||  
QY 766 ATCTCTGTACGGGATGTGCGGCGCGGGCTTCTGTTTCAACACCTATCCG 825  
Db |||||  
QY 781 GCGCAGGTCTTTCATGCGCGAGCTGCGCGCTGCGCTGGCGCGCGCTGGGACCATC 840  
Db |||||  
QY 826 GCGCAAGTCTTTATGCGCGAGCTGCGTTCATCTGCGCTCGGCGCGGCTTGGGCAATATC 885  
Db |||||  
QY 841 GCGGTATGTCGCCAGGAGATCGTCTGTTTATCATGAGTGGTGGGTGTTGTCATCGAA 900  
Db |||||  
QY 886 GCGGTGCTGCTGCGTACAGGATTCCTGCTGTTGATCATGCGCGGGGTTTTGCGTGGTGA 945  
Db |||||  
QY 901 ACCCTCTCGGTGATGATCCAGGTGCTTCTTCAAGCTG-ACCGGACGCGCGCTTCCG 959  
Db |||||  
QY 946 ACCTGTGCTGTTTCTGCGAGTGGCTCTTTAAAGCTGCGCGGTGAGCGCATCTCCG 1005  
Db |||||  
QY 960 TATGCGCGCGATCCATCACCATTTCGA 986  
Db |||||  
QY 1006 CATGCGCGCGATTCCACCACCTATGA 1032  
Db |||||

## RESULT 4

US-09-543-681A-2879  
; Sequence 2879, Application US/09543681A  
; Patent No. 6605709  
; GENERAL INFORMATION:  
; APPLICANT: GARY BRETON  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS  
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 2709.1002-001  
; CURRENT APPLICATION NUMBER: US/09/543,681A  
; PRIOR FILING DATE: 2000-04-05  
; PRIOR APPLICATION NUMBER: US 60/128,706  
; PRIOR FILING DATE: 1999-04-09  
; NUMBER OF SEQ ID NOS: 8344  
; SEQ ID NO 2879  
; LENGTH: 1083  
; TYPE: DNA  
; ORGANISM: Proteus mirabilis  
US-09-543-681A-2879

Query Match 33.8%; Score 366.2; DB 4; Length 1083;  
Best Local Similarity 58.6%; Pred. No. 7e-74;  
Matches 635; Conservative 0; Mismatches 448; Indels 0; Gaps 0;

QY 1 ATGCTCCTGTGCTGGCGGAATACCTGCAACAGTTCTTACAAGGCTTCGGCGTTCAG 60  
Db |||||  
QY 61 TACCTGACCTCGCGGCAATCTCAGCGTCTCAGCGGTGTCGCTGCTGCTGGCTG 120  
Db |||||  
QY 61 TATCTGACATTCAGAGCGATTGTTGGTTTATGACGCAATTAATTATGCGTTATGATG 120  
Db |||||  
QY 121 GGGCCCTGGATGATCCGTACCTTCAGATCCCCAGATCGGCCAGGCGCTGCGCAACGAC 180  
Db |||||  
QY 181 GGTGCGAGTGCACCTGTGGAAGAGGCAACCCGACCATGGCGGGCGCTGATCCTT 240  
Db |||||  
QY 181 GCGCTGTAGTCACACTTTAGCAACAGTGTACGCGGACCATGGGTGGGATCATGATCCTA 240  
Db |||||  
QY 241 ACCGCGATAGCAGCAGCAGCTCTGTGGGCGGATCTTTCCAAACCGCTACGTGTGGGTA 300  
Db |||||  
QY 241 TTTCTATCGAGTATCAACCTTATTGTGGGCGGATAGATATCTTATGATGTGT 300  
Db |||||

QY 301 GTGCTGGTCTTACCGTCTGTCGTCGTCATCGCTGGGTAGACGACTACCGCAAGTG 360  
 Db 301 GTACTCTTGGTCTTAAATGGTTATGGCATCATCGTTTTATCGATGATATCGCAAGTA 360  
 QY 361 ATCGAGAAAGAACTCCCGTGGCTCCGAGCGCTGGAGTACTTCTGGCAGTGGGTTC 420  
 Db 361 GTGGTAAAGATATCTCAGGCTTAAATAGCGGTTGGAATATTTCTGGCAATCGGTATTA 420  
 QY 421 GGCATCGCGCGCGGTGTTCTCTACATGACTCCGAAACCCGATCGAGACACCGCTG 480  
 Db 421 GCTTTAGCGGTGATTTAGTATGTAACGATGTAAGATACCCAGCAAGCAACTT 480  
 QY 481 ATCGTGGCGATGCTGAAGAGCGTCCAGATCCAGTTGGGATCTTCTTCTGGTCTGAC 540  
 Db 481 GTCGTGGCTTTCTTTAAAGATGTATGCCCAACTTGGCATGTTGTATATCTTACTGGCT 540  
 QY 541 TACTTCGTCATCGTCGCTCCAGCAATCGGTGAACTCAGCAAGCTCTCGACGCGCTG 600  
 Db 541 TATTTTGTCATGTTGGTACCAAGTACCGGTAAATTTAACCGATGTTTAGATGGCTTA 600  
 QY 601 CGCATATCGCGCGGTAAATGTTCCGCGCGCTGGGCATCTTCTGCTACCTGTCCGGC 660  
 Db 601 GCCATTATCGCTCGCTGTTGTAGCGCTGGATTTGCATTAGTTCATGGCTACAGT 660  
 QY 661 AACGTGAAGTTCCGCGAGTACTGCTGATTTCCCAAGTACCGGCGCGCGGAGCTGATC 720  
 Db 661 AATGTCAATTTTGGCAGCTACTAAAATTCCTTTAATGATGCAATGCGCGAATTGGTG 720  
 QY 721 GTGTTCTCGCGCGCTGGTCCGCGCGCTCGGCTTCTCTGTTTCAACACCTATCCG 780  
 Db 721 ATTGTTGTACCGCCATTGTCGGTGGGATTAGCTTCTTGGTTTAACTATCTTATCCT 780  
 QY 781 GCGCAGTCTTCAATGGGCGAGTCCGCGCGCTGGCGCTGGCGCGCGCTGGCGCAACATC 840  
 Db 781 GCCCAAGTCATTATGGGTGATGTTGGCTCTTTAGCATTAGTGGCGCATTAGTACTATC 840  
 QY 841 CGGTGTATCGTCCGCGAGAGATCGTGTGTTTCATCATGTTGGTGGGTGTTTCGTCATGAA 900  
 Db 841 GCGGTGTTATACGCCAAGATTTTATTAGTGAATATGGGCGGTGTTTCGTTGTGAA 900  
 QY 901 ACCCTCTGGTATGATCCAGTCTGCTTCTTCAAGTCTACCGAGCGCGCGCTTCCGT 960  
 Db 901 ACGTCTCGGTATTTTAAAGTAGTGTTCATTTAACTGCGCGTCAACGTATTTCCGT 960  
 QY 961 ATGGCGCGATCCATCACATTTGAACTGAAAGCTGGCGGACCGCGCTGATCGTG 1020  
 Db 961 ATGGCGCAATCCATCATCATATGAATTAAGGTTGGCCAGAGCTCTGTGTGATGTC 1020  
 QY 1021 CGCTTCTGGATCATCACGCTGATCTGTTGCTGATGCGCTCGCCACCTTGAAGCTCGT 1080  
 Db 1021 CGCTTTGGATCATCTCTTATGTTAGTCTTATGTTAGTCTTATGTTAGCGACATTAAGGTACGT 1080  
 QY 1081 TGA 1083  
 Db 1081 TAA 1083

RESULT 5

US-09-557-884-1  
 ; Sequence 1, Application US/09557884  
 ; Patent No. 6506581  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Fleischmann et al.  
 ; TITLE OF INVENTION: The Nucleotide sequence of  
 ; the Haemophilus influenzae Rd Genome, Fragments  
 ; Thereof, and Uses thereof  
 ; NUMBER OF SEQUENCES: 1  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Human Genome Sciences, Inc.  
 ; STREET: 9410 Key West Avenue  
 ; CITY: Rockville  
 ; STATE: MD

COUNTRY: USA

ZIP: 20850

COMPUTER READABLE FORM:

MEDIUM TYPE: 3 1/2 inch diskette

COMPUTER: Dell Pentium

OPERATING SYSTEM: MS DOS v6.22

SOFTWARE: ASCII Text

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/557,884

FILING DATE: 25-Apr-2000

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/476,102

FILING DATE: JUN-5-1995

ATTORNEY/AGENT INFORMATION:

NAME: Michelle S. Marks

REGISTRATION NUMBER: 41,971

REFERENCE/DOCKET NUMBER: PB186P3

TELECOMMUNICATION INFORMATION:

TELEPHONE: 301-309-8504

TELEFAX: 301-309-8439

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 1830121 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 1:

US-09-557-884-1

Query Match 33.4%; Score 361.4; DB 4; Length 1830121;

Best Local Similarity 58.4%; Pred. No. 4.6e-72;

Matches 632; Conservative 0; Mismatches 451; Indels 0; Gaps 0;

QY 1 ATGCTCTGCTGCTGGCGAATACCTGCAAGTTCTACAAGGGCTTCGGCGTCTTCAG 60  
 Db 1201927 ATGTTAGTCTGCTGCTGAATATCTTGTCTGTTACGAAACCGCGCTTTAATGCTATTCT 1201986  
 QY 61 TACCTGACCTCGCGGATCTCTACGCTGCTCACCGCGCTGTCGCTGCTGCTGGCTG 120  
 Db 1201987 TATTTACGTCGCGGCAAACTTCGATTTAAACCGACATTTTATCTCATTTGGATT 1202046  
 QY 121 GGGCCCTGGAGATCCGTACCTTGAGATCCCCAGATCGCCAGCGCGTGGCAACGAC 180  
 Db 1202047 GGCCTTAAAGTGATCAACACGCTTGCAATCTTAAATTTGGCAAGAGTGCATATGAT 1202106  
 QY 181 GGTCCGAGTCCGACCTGTCGAGAGAGGCAACCGCATGGCGCGCGCTGATCCTT 240  
 Db 1202107 GGCCTTGAAGTCACTTTGCAAAAAGGACACCCACTATGGGTGGTGTGATTTTA 1202166  
 QY 241 ACCGCCATAGCCATCAGCACGCTGCTGTGGCGGATCTTTCCAAACCGCTACGTGGGTA 300  
 Db 1202167 TTCTCTATTGGCGTAAGTAGCTTATTATGGGCAATCTTGCTAATCCGTATATTTGGGTT 1202226  
 QY 301 GTGCTGGTCTTACCTGCTGTTCCGTGCCATCGCTGGGTAGACGACTACCGCAGGTG 360  
 Db 1202227 TGTATTATTGTTTATTATTTGGATACCGCGCAATTTGTTGGATGATTTCCGTAATAATT 1202286  
 QY 361 ATCGAGAAAGAACTCCCGTGGCTGCGGAGCGCTGCGAAGTACTTCTGCGAGTCGGGTGTC 420  
 Db 1202287 ACCGTAAAAATACTGATGATTGATTGCTCGTTGGAATAATTTCTGGATGCTGTGGTG 1202346  
 QY 421 GGCATCGGCGCGCGTGTTCCTCTACATGACTCGCGAAACCCCGATCGAGACACCGCTG 480  
 Db 1202347 GCATTAGTGGCAATCCTTTGGCTTTTATTGGCTGTGTCAGCACACTGATGCCACCGCTTA 1202406  
 QY 481 ATCGTGGCGATGCTGAAGAGCGTCCAGATCCAGTTGGGATCTTCTTGGTCTCGTACC 540  
 Db 1202407 GTGATTCATTCTTTAAGACATTTATGCTCAATTAGGTTGTTCTATATTGTTTACT 1202466  
 QY 541 TACTTGTGTCATCGTGGGTCCGAGCAATCGGTGAACTCAACGACGGTCTCGACCGGCTG 600  
 Db 1202467 TACTTTGTGATTGTTGGCAGCGGTAAATGCAAGTGAATTTAAACCGAGTTTAGATGATTA 1202526

601 GCATCATGCGCGGTAATGTTGCGGCGCGCTGGCGATCTTCTGCTACCTGTGCGGC 660  
1202527 GCATATGCTACTGCGCTGTTGAGGTCGCTTTCCTTATTCTTGGCTACAGT 1202586  
661 AAGTGAAGTTCGCGAGTACCTGCTGATTCACAGCTACCGGCGCGCGAGCTGATC 720  
1202587 AAGTGAATTCGAGAATATTTGCATATCCGATATTAATTAATACAGTTCTGAAGTAGTG 1202646  
721 GTGTTCTGCGCGCGCTGTTGCGGCGCGCTGCGCTTCTGCTGCTTCAACACCTATCG 780  
1202647 GTGTTCTGACAGTAATGTTGCGGAGTTTGGGATCTTAATGTTTAATCTATCCA 1202706  
781 GCGAGCTCTTCATGCGCGAGCTGCGGCGCGCTGCGCTGCGGCGCGCTGCGGCGCATC 840  
1202707 GCTCAAGTATTTATGCGGAGCTGCGGCTTCTAGCATTTAGTGTGCTGCTGTTGTA 1202766  
841 GCGGTGATCTGCGCGAGGAGCTGCTGCTTCAATGCTGCGGCTGCGGCGCGCTGCGGCGAA 900  
1202767 GCAATCTGTTTCTGTCAGGAATTTTTCCTGCTGATTAATGCGTGTGATTTGTTGTA 1202826  
901 ACCCTCTCGGTGATGCCAGGTGCTTCTTCAAGCTGACCGGAGCGCGCGCTTCCGT 960  
1202827 GCACTCTCTGTTATTTGCAAGTAGGCTCTTAAGTTACGCAACAAACGCAATTTTGA 1202886  
961 ATGCGCGGATCCATCACCAATTTGCAACTGAAGGCTGCGCGGAGCGCGCGCTGATCGTG 1020  
1202887 ATGCGCGGATCCACCAATTTGCAATGAAGGATGCGCTGAGCGGAGGATTAAT 1202946  
1021 CGTTCTGATCATCACCGTATCCTGCTGATCGGCTCGCCACCTTGAAGCTGCGT 1080  
1202947 CGGTTTGGATTAATTTCTTAATGCTGCTGCTGATGGAATGCTCACCTTGAAGTTGCGT 1203006  
1081 TGA 1083  
1203007 TAA 1203009

RESULT 6  
US-09-643-990A-1  
Sequence 1, Application US/09643990A  
Patent No. 6528289  
GENERAL INFORMATION:  
APPLICANT: Robert D. Fleischmann  
Mark D. Adams  
Owen White  
Hamilton O. Smith  
J. Craig Venter  
TITLE OF INVENTION: The Nucleotide sequence of  
the Haemophilus influenzae Rd Genome, Fragments  
Thereof, and Uses Thereof  
NUMBER OF SEQUENCES: 1  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Human Genome Sciences, Inc.  
STREET: 9410 Key West Avenue  
CITY: Rockville,  
STATE: MD  
COUNTRY: USA  
ZIP: 20850  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3 1/2 inch diskette  
COMPUTER: Dell Pentium  
OPERATING SYSTEM: MS DOS v6.22  
SOFTWARE: ASCII Text  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/643,990A  
FILING DATE: 23-Aug-2000  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/487,429  
FILING DATE: 1995-06-07  
APPLICATION NUMBER: 08/426,787  
FILING DATE: 1995-04-21

ATTORNEY/AGENT INFORMATION:  
NAME: Kenley K. Hoover  
REGISTRATION NUMBER: 40,302  
REFERENCE/DOCKET NUMBER: PBI86P1C1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 301-610-5790  
TELEFAX: 310-309-8439  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1830121 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
US-09-643-990A-1  
Query Match 33.4%; Score 361.4; DB 4; Length 1830121;  
Best Local Similarity 58.4%; Pred. No. 4.6e-72;  
Matches 632; Conservative 0; Mismatches 451; Indels 0; Gaps 0;  
QY 1 ATGCTCTGCTGCGCGGAAATACCTGCAACAGTTCTACAGGCGCTTCGGCGCTTCCAG 60  
Db 1201927 ATGTTAGTCTGCTGCTGAAATATCTTGTCTTACGAAACCGGCTTAATGCTATTCT 1201986  
QY 61 TACCTGACCTGCGCGGCAATTCAGCGTCTCACCGGCTGCTGCTGCTGCTGCGCTG 120  
Db 1201987 TATATTACGTCGCGGCAATTCCTTGCATTATTAACCGCACTTTTATCTCATTTGGATT 1202045  
QY 121 GGGCCCTGGATCCGTACCTTCAGATCCCGCAGATCGGCCAGGCGCTGCGCAACGAC 180  
Db 1202047 GGGCCCTAAAGTATCAACAGCTTCAGATCTTAAATTTGCGCAAGAGTGCAGAAATGAT 1202106  
QY 181 GGTCCGAGTCGCACTGTCGAGAGGCGCACCCGACCATGCGGCGGCGGCGCTGATCCTT 240  
Db 1202107 GGGCCCTGAAAGTCACTTTGCAAAAAGAGCACCCCATATGGGTGGTGTGATGATTTTA 1202166  
QY 241 ACCGCCATAGCCATCAGCACGCTGCTGTGGCGGATCTTTCCAAACCGCTACGTGGGTA 300  
Db 1202167 TTCTCTATTGGCGTAAGTACGTTATTATGGCAAAATCTTGCTAATCCGTATATTGGGTT 1202226  
QY 301 GTGCTGTGCTTACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360  
Db 1202227 TGTTTATTGTTTATTGTTGATACGCGCAATGTTGTTGTTGTTGTTGTTGTTGTT 1202286  
QY 361 ATCGAAGAACTCCCGTGGCTGCGGAGCGCTGGAAGTACTTCTGCGAGTCTGCTGCTGCTG 420  
Db 1202287 ACCCGTAAATATCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1202346  
QY 421 GGCATCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480  
Db 1202347 GCATTAGTGGCAATCTTTGGCTTTTATTTGGCTTGTCTACGACACTGATGCCACCCGTTTA 1202406  
QY 481 ATCGTGGCGGATGCTGAAGAGGCTCGAGATCCAGTTGGGCATCTTCTTCTGCTGCTGCTG 540  
Db 1202407 GTGATTCATCTTTAAAGACATTAATGCTCAATAGGTTTGTCTATATTTGTTTATCT 1202466  
QY 541 TACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600  
Db 1202467 TACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1202526  
QY 601 GCGATCATGCCGAGTAATGTTGCGCGCGCTGCGGCTGCGGCTGCTGCTGCTGCTGCTG 660  
Db 1202527 GCGATTATGCTTACTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1202586  
QY 661 AACGTGAAGTTCCGCGAGTACCTGCTGATTTCCCAACGTAACCGGCGCGCGCGGAGCTGATC 720  
Db 1202587 AACGTGAATTTGCGAGAATATTTGCAATTCGATATTAATTAACAGTTCTGAAGTAGTG 1202646  
QY 721 GTGTTCTGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780  
Db 1202647 GTGTTCTGCTACAGCTATTTGTTGCGCGAGTTTGGGATTTCTTATGTTTAACTACTTATCCA 1202706  
QY 781 GCGCAGGTTCTTCATGGCGGAGCTCGCGGCTGCGGCTGCGGCGCGCGCGCGCGCGCGCGC 840

Db 1202707 GCTCAAGTATTTATGGGGGAGCGTGGTCTCTAGCAATAGTGTGGCTGGGTGTGTA 1202766  
Qy 841 GGGGTGATCGTGGCCGACGAGATCGTGTTCATCATGATGGTGGGGTTCGTCATGGAA 900  
Db 1202767 GCAATCTTGTGTGTCAGGAATTTTGGCTGTGATTAATGGTGGTGTATTTGTGTTGAA 1202826  
Qy 901 ACCCTCTCGGTGATGATCCAGTCCGCTTCTCAAGCTGACCGGACGGCGGCTTCGCT 960  
Db 1202827 GCATCTCTGTATTTTGGAGTAGGTCCTTATAGTTACGAAACACGATTTTAGA 1202886  
Qy 961 ATGGCCGATTCATCACATTTGAACTGAAAGCTGGCCGACCCGCGCGTGAATCGTG 1020  
Db 1202887 ATGGCACCGCATTCACCAACATTTTGAATGAAGGATGGCTGAGCCCAAGAGTGAATTT 1202946  
Qy 1021 CGCTCTCGATCATCACCGTGAATCGTGTGCTGATCGGCCCTCGCACCTTCAAGCTGCGT 1080  
Db 1202947 CGGTTTGGATTAATTTCTTAAATGCTGTGTGATGGGATTTGGTCACTTGAAGTTGCGT 1203006  
Qy 1081 TGA 1083  
Db 1203007 TAA 1203009

## RESULT 7

US-09-328-352-2043  
; Sequence 2043, Application US/09328352  
; Patent No. 6562958  
; GENERAL INFORMATION:  
; APPLICANT: Gary L. Breton et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER  
; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: GTC99-03PA  
; CURRENT APPLICATION NUMBER: US/09/328,352  
; CURRENT FILING DATE: 1999-06-04  
; NUMBER OF SEQ ID NOS: 8252  
; SEQ ID NO 2043  
; LENGTH: 1119  
; TYPE: DNA  
; ORGANISM: Acinetobacter baumannii.  
US-09-328-352-2043

Query Match 23.1%; Score 250.6; DB 4; Length 1119;  
Best Local Similarity 53.4%; Pred. No. 8.7e-48;  
Matches 597; Conservative 0; Mismatches 484; Indels 36; Gaps 2;  
Qy 1 ATGCTCTGCTGTCGCGGCAATACCTGCAACAGTCTACAGGGCTTCGGGCTTTCCAG 60  
Db 1 ATGCTGTTATGGTTGTTTGAACAACCTTGGCGGCTATCACAGTTCGTTTCAGGTTGTTGCT 60  
Qy 61 TACCTGACCTCGCGGGCAATCTCAGCGTGTCTACCGGCTGCTGCTGCTGCTGCTGCTG 120  
Db 61 TATTTAAGATTAGTTCCTTACTCAGTGTATTAAGTCACTGACCACTGGTTGGTTCTC 120  
Qy 121 GGGCCCTGGATGATCGTACCTGTCAGATCCCGCAGATCGGCCAGCGCGTGGCAACAC 180  
Db 121 GGACCAATCATGATTCGTAATTAACAAGGTTAAATACCGTTCAGGCGATGAATTCGTT 180  
Qy 181 GGTCCGACGTCGACCTGTCGAGAGGCGACCCGACCATGGCGGCGCTGATCCTT 240  
Db 181 GCTCTGAAATCATGCTAAGAGATGGGTACCGACCATGGTGGGATTTAATTCG 240  
Qy 241 ACCGCCATGCCATCAGCAGCTGCTGTGGCGGATCTTTCCACCGCTAGCTGTGGTA 300  
Db 241 CTCTCAATTGGTATTAGTACTTTATTGGGCTGATTTATCTAAACCTTATGTTGGATT 300  
Qy 301 GTGCTGGTGTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360  
Db 301 GTACTTGGTGTATGTTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360  
Qy 361 ATCGAGAAGAACTCCCGTGGCTGCGGACCGCTGGAAGTACTTCTGGCAGTCGCTGCTC 420  
Db 361 CGCTATAAAGATAATGACAGTCTACCTGCGCGTGAAGAGTCTTCTGGACTTCTGCGCA 420

Qy 421 GGCATCGGCGCGCGTGTCTCTACATGACTG---CCGAAACCCGATCGAGACCAC 476  
Db 421 TCGCTGGGTGCTGGTATTCGGTTATATTTAAATGCTACCCAAATCTAATCGAATAC 480  
Qy 477 CTTGATCGTGGCGATGCTGAAGAGCGTCCAG----- 507  
Db 481 ACAGCAAAATATGCTGATTTATGATTCCTTTTAAAGAACTCTTCAATCCGCTTCT 540  
Qy 508 ---ATCCAGTTGGGATCTTCTCGTGTCTGACTACTTCTGATCGTCCGCTCGAGC 564  
Db 541 ATGTTCTCTTAGTCTTAGCATTTATGTAATTTACTATTTAGTAATTAATGTCCTCT 600  
Qy 565 AATGCGGTGAACCTCACCGACGCTCTCGACGCGCTGGCGATCATCGCGAGTAATGTT 624  
Db 601 AATGCGGTTAACTTAATCTAGTCTAGATGGTTAGCCATTAATGCGAGTAGTATGTA 660  
Qy 625 GCGGCGCGCTGGGATCTTCTGCTACCTGTCGGGCAAGTGAACTTCGCGAGTACCTG 684  
Db 661 GCGACGGGCTTAGGCGTATTTGCGTATTTATCTGGTATTCGTTTTCGCAACTATTG 720  
Qy 685 CTGATTTCCCAACGTAACCGGCGCGCGAGCTGATCTGTTCTGGCGCGCGCTGGTCCGC 744  
Db 721 CATATTCCTTATGTGAATACACTTCAGAACTTGGTGATCTGTTCTGCCATGATTGT 780  
Qy 745 GCGGCGCTCGGCTTCTCTGTTTCAACCTATCGGCGCAGTCTTCTATGGCGACGTC 804  
Db 781 GCGGCTTTGGCATTTCTTTGGTATAACGCACATCGGCGCAAGTCTTTATGGTGATGTG 840  
Qy 805 GCGGCGCTGGCGTGGCGCGCGCTGGGCACCATCGCGTGTATCGTGGCGCGAGATC 864  
Db 841 GCGGCTTTAGCTAGTGTGCAATGCTTGGAAACCATTCGCTTATGTTGTCGCCAAGAAC 900  
Qy 865 GTGCTGTTCAATCANGGTGGGTGTTGCTGATGAGAAACCTCTCGGTGATGATCCAGTTC 924  
Db 901 GTATTCGCAATTAATGGCGGTGTTTTCGTTATGGAAGCGGTATCGGTATTCCTGCAAAATC 960  
Qy 925 GCTTCTCTCAAGCTGACCGGCGCGCTCTCCGCTATGGCGCGATCCATCACCACTTC 984  
Db 961 GGCTCATAGGATGGGATTAACCGTATTTTATGGCACCTTTGCAACCATCACTAT 1020  
Qy 985 GAATGAAAGGTGGCGGACCGCGCTGATGCTGCGCTTCTGATCATCACCCTGATC 1044  
Db 1021 GAAACCAAGGCTGGAAGAAACCCAAAGTAGTATCCGTTCTGGATTTATGATGATTG 1080  
Qy 1045 CTGCTGCTGATCGCTCGCCACCTTGAAGCTGCGTT 1081  
Db 1081 CTGCTGTTTATAGGCTTAATGACCTTAAATTCGCT 1117

## RESULT 8

US-09-540-236-1218  
; Sequence 1218, Application US/09540236  
; Patent No. 8673910  
; GENERAL INFORMATION:  
; APPLICANT: Gary L. Breton et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATAI  
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 2709.2005-001  
; CURRENT APPLICATION NUMBER: US/09/540,236  
; CURRENT FILING DATE: 2000-04-04  
; NUMBER OF SEQ ID NOS: 3840  
; SEQ ID NO 1218  
; LENGTH: 1107  
; TYPE: DNA  
; ORGANISM: M. catarrhalis  
US-09-540-236-1218

Query Match 21.6%; Score 233.8; DB 4; Length 1107;  
Best Local Similarity 52.6%; Pred. No. 5.4e-44;  
Matches 560; Conservative 0; Mismatches 472; Indels 33; Gaps 1;  
Qy 52 GTCTTCCAGTACCTGACCCCTGCGCGCATTTCTCAGCGTGTCTACCGGCTGCTGCTGCTG 111

Db 43 GCATCGGTTTCAATGACTTTTGGCGGCACTGCTTGGCGGTGATTACATCGCTGCTATTGTT 102  
Qy 112 CTGTGGCTGGGCGCTGGATGATCCGTACCTTCAGATCCCGAGATCCCGAGCGCGTG 171  
Db 103 ATCAATGGTGAAGCTGTGATTCATATCTGCGTACCTTAAATACCGTCAAGCGGTG 162  
Qy 172 CGCAAGACGGTCCGAGTCCGACCTGTGGAAGAGGCAACCCGACCACTGGCGCGCC 231  
Db 163 CGTGATGATGAGCAAAATCACTACCTTGCCAAACAGGACGCTACGATGGCGGGGTG 222  
Qy 232 CTGATCCTTACCGCATAGCATCAGACGCTGCTGGGCGGATCTTCCCAACGCTAC 291  
Db 223 CTGATTTTGGTGGGATGGTATGCTACCTTGGCATGGCGGATTTAGCAATCCTTAT 282  
Qy 292 GTGTGGTGTGCTGCTGCTACCTGCTGCTTGGTGGCCATCGCTGGGTAGACGATAC 351  
Db 283 GTTTGGATTTTAATGGTGGTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 342  
Qy 352 CGCAAGGTGATCAGAGACTCCGCTGGCTGCTGGCGGCTGGAGTACTTCTGGCAG 411  
Db 343 CTAAATCAAGCAAAACCCCAAGGACTGATGCTCGCAAAATATTTTGGCTA 402  
Qy 412 TCGGTGTTCGGCATCGCGCGCGCTGCTCTCATGACTGCGGAAACCCCGCATCGAG 471  
Db 403 TCGGTGCGCTCGCTATTGCTGGCTGTTCGATGTATTACATGCCACAGCAAGATACC 462  
Qy 472 ACCACCTGATCGTGGCGATGCTGAAGAGCTC----- 504  
Db 463 AATACCATGATCGCCATGCAAGACATGCTAATCCACTGTTTAAAGACATTTGATTCOG 522  
Qy 505 -----GAGATCCAGTTGGGCACTTCTTGGTCTGCTGACCTACCTTCGTCATCGTGGCG 558  
Db 523 TTTTCAGCATCCATTTGGGTTTGGGATTTATTTATTCGCACTTATTTTGTGTTGGCAGC 582  
Qy 559 TCAGCAATCGGTGAACCTCACGAGCGTCTCGACGGCTGGCGATCATTCGCGACGGTA 618  
Db 583 TCATCTAATCGGTGAATCTGACAGATGGTTGGATGGATTTGATTTTACCCGTAGTT 642  
Qy 619 ATGGTTGCGCGCGCTGGGCACTTCTTCTGCTACTGTCGGGCAACGTGAAGTTGCGCGAG 678  
Db 643 TTGGTGCAGCAGGCTTGGGCGTTTTCCTTATTTCTGGCTGCGCAATTTGCTGAT 702  
Qy 679 TACCTGCTGATTCCTCAACGATACCGGGCGCGGAGCTGATGCTGTTCTGCGCGCGCTG 738  
Db 703 TATATGATGCTGCTTATATTGCTTATATGCTGAGGTTACCATCGTCTGTGCTGCCATG 762  
Qy 739 GTGGCGCGCGCTCGCTTCTCTGTTCAACACCTATCCGCGCGAGGCTTCTATGGCG 798  
Db 763 ATTGGCGCAGGCTTGGGCTTCTTGGTATAATGACAGACCCCGCATGTATTTATGGGC 822  
Qy 799 GAGTCTGGCGGCTGGGCGCTGGGCGCGCTGGGCAACCATCGCGGTGATCGTGGCCAG 858  
Db 823 GATCTGGGTGATTTGAGTCTTGGGGGTATGCTTTGGTACCATTGCTGTGATGACTCGCCAA 882  
Qy 859 GAGATCTGTGTTTCATCATGGTGGGTTGCTGATGGAACCCCTCTCGGTGATGATC 918  
Db 883 GAGTTGGCGTTGCGATCATGGGCGGTATCTTTGTGGCTGAAACGCTATCGGTGATTTTA 942  
Qy 919 CAGTCTGCTTCTTCAAGCTGACGGACGCGCTTCTCGTATGCGCGCATCCATCAC 978  
Db 943 CAGGTTGTTCTTATCGTCTGGTAAACACGCAATTTTGTGATGGCACTTTACATCAT 1002  
Qy 979 CATTTGAACTGAAGCTGGCGGACCCCGCGGTGATGCTGGCTTCTGATCATCAC 1038  
Db 1003 CATTTTGAAGAGCTTGGACTCAAGAAACCAAAAGTAGTTCGCGGCTTTTATATTGTGT 1062  
Qy 1039 GTGATCTGTGCTGATCGGCTCGCCACCTTGAAGCTGGTTGA 1083  
Db 1063 ATTAATTTTGGTGGCTTGGACTGATGACCTTAAATACGCTAA 1107

RESULT 9

US-09-596-002-39  
; Sequence 39, Application US/09596002  
; Patent No. 6632636  
; GENERAL INFORMATION:  
; APPLICANT: Lagace, Robert, E.  
; APPLICANT: Patterson, Chandra  
; APPLICANT: Berg, Kim, L.  
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES OF MORAXELLA CATARRHALIS GENOME  
; FILE REFERENCE: PM-0008-4 US  
; CURRENT APPLICATION NUMBER: US/09/596.002  
; CURRENT FILING DATE: 2000-06-16  
; PRIOR APPLICATION NUMBER: 60/140,121  
; PRIOR FILING DATE: 1999-06-18  
; NUMBER OF SEQ ID NOS: 41  
; SOFTWARE: PERL Program  
; SEQ ID NO 39  
; LENGTH: 100848  
; TYPE: DNA  
; ORGANISM: M. catarrhalis  
; FEATURE:  
; NAME/KEY: misc\_feature  
; OTHER INFORMATION: Incyte template ID No. 6632636 39  
; PUBLICATION INFORMATION:  
US-09-596-002-39

Query Match 21.4%; Score 232.2; DB 4; Length 100848;  
Best Local Similarity 52.5%; Pred. No. 3.4e-43;  
Matches 559; Conservative 0; Mismatches 473; Indels 33; Gaps 1;  
Qy 52 GTCCTCCAGTACCTGACCCCTGCGCGGATTCCTCAGCTGTCTCAGCGCTGTCTCCTGTGCG 111  
Db 76209 GCCATCGGTTTCATGACTTTTGGCGGCACTGCTTGGCGGTGATTACATCGCTCTTATTGTT 76268  
Qy 112 CTGTGGCTGGGCGCTGGATGATCCGTACCTTTCAGATCCCGAGATCCCGAGATCGGCGCGCTG 171  
Db 76269 ATCATTGGTGAAGCCCTGTGATTCATTTCTGCTGCTTAAATACGCTCAAGCCGCTG 76328  
Qy 172 CGCAACGACGGTCCGCACTGCGACCTGTGGAAGAGGCGACCCCGACCATGGCGCGGCC 231  
Db 76329 CGTGATGATGACCAAAATCACTTGGCCAAACAGGACGCTACGATGGCGCGGTG 76388  
Qy 232 CTGATCTTACCGCATAGCCATCAGACGCTGTGTTGGGCGGATCTTCCAAACGCTAC 291  
Db 76389 CTGATTTTGGTGGCGATTTGCTTATGCTACCTTGGCATGGGCGGATTTAGGCAATCCTTAT 76448  
Qy 292 GTGTGGGTAGTGTGCTGTTTACCTGCTGTTCGGTGCCATCGGCTGGGTAGACGATAC 351  
Db 76449 GTTTGGATTTTAAATGGTGGTTCATGCTGCTATTTTGGTGCAGTTGGCTGGGAGATGATGG 76508  
Qy 352 CGCAAGGTGATCGAAGAACTCCCGTGGCTGCGGAGCCGCTGGAAGTACTTCTTGGCAG 411  
Db 76509 CTAAAAATCAAGCAAAACCCCAAGGACTGATTTGCTGCAAAAAATATTTTGGCTA 76568  
Qy 412 TCGGTGTTTCGGCATCGGCGCGCGCTGTTCTCTACATGCTGCCGAAACCCCGATCGAG 471  
Db 76569 TCGGTGCGGCTCGCTATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 76628  
Qy 472 ACCACCTGATCGTGGCGGATGCTGAAAGAGCGTC----- 504  
Db 76629 ACTGCCATGATCGCCATGCAAGACATGCTAATCCACTGTTTAAAGACATTTGATTCG 76688  
Qy 505 -----GAGATCCAGTTGGGCACTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 558  
Db 76689 TTTTCAGCGATTCATTTGGGTTTGGGATTTAATTTGCGACTTATTTTGTGTTGGCAGGC 76748  
Qy 559 TCGAGCAATGGGTGAACCTTCACGACGCTTCGACGCGCTGGGATCATGCCGACGCTA 618  
Db 76749 TCATCTAATGCGGTGAATCTCAGAGATGGTTTGGATGGATTTGATTTTACCCGCTAGTT 76808  
Qy 619 ATGGTTGCGGCGCGCTGGGCACTTCTGCTACCTGCTCGGCAACGATGAGCTTCCGCGAG 678  
Db 76809 TTGGTCCGACGAGGCTTGGGCGTTTGTGCTTATATTTTGGCTCGCCACATTTTGTCTGAT 76868





Db 241823 GTTGGCACTAGTAATCAGTTAATCTAACAGATGGAATAGATTGTTTATGCTATTATGCCA 241764  
Qy 613 ACAGTAATGTTGCGCGCGCTGCGGCACTCTTCTGCTACCTGTCGCGGCAACGTGAAGTTC 672  
Db 241763 GTTATTTTAACTTGTGTTGACATTAATCTTTATTTAGCGACAATAATAATATC 241704  
Qy 673 GCGAGTACCTGCTGATCCCAAGTACCGGCGCGCGGAGTGTGCTGCTGCGGC 732  
Db 241703 TCACATTATCTACATGTTTAAATATGTAATAAATCAACAGACTAGCTATATATG 241644  
Qy 733 GCGTGTGCGCGCGCGCTGCGCTCTCTGCTCAACACCTATCCGCGGAGGTCTTC 792  
Db 241643 GCAATGTTGTTGAGATAGTTTATGTTTAAAGCTATCCAGCAAAAGTTC 241584  
Qy 793 ATGGGCACTGCGCGCGCTGCGCGCTGCGCGCGCTGCGCGCGCTGCGCGCGCTGCG 852  
Db 241583 ATGGGAGATGTTGATCACTAGCATAGTGGTCTTTTAGGAGCTATAGCAATATTATTA 241524  
Qy 853 GCGCAGGAGATCGTCTGTTTCATCATGGTGGGTGTTGCTCATGAAACCTCTCGGTG 912  
Db 241523 CATCAAGATTAATCAATTAATCAATGAGTGGTATTTTGTGTTGAACAATATCTG 241464  
Qy 913 ATGATCCAGTCTGCTCTCTCAAGCTGACCGGCGCGCGCTCTTCTGATGGCGCGATC 972  
Db 241463 ATATACAAATATATCGTTTAAATTAAGAAAAAAGAAATTTTCAATGGCACCAGT 241404  
Qy 973 CATCACCATTTCGAAGTGGCGCGCGCGCGCGCGCGCTGCTGCGCTCTTCTGATC 1032  
Db 241403 CACCACCAATATGAGTCAAGGAATATGAACTTGTATCTGTAAGATTGATTTG 241344  
Qy 1033 ATCACCCTGATCTGCTGATCGGCTCGCCACCTTGAGCTGCTTGA 1083  
Db 241343 GTGTCACATAATGTTACTATAAGTCTCATTTCTTAAAGGTATGTTGA 241293

RESULT 13  
US-09-489-039A-5281  
; Sequence 5281, Application US/09489039A  
; Patent No. 6610836  
; GENERAL INFORMATION:  
; APPLICANT: Gary Breton et. al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA  
; FILE REFERENCE: 2709.2004001  
; CURRENT APPLICATION NUMBER: US/09/489,039A  
; PRIOR FILING DATE: 2000-01-27  
; PRIOR FILING DATE: 1999-01-29  
; NUMBER OF SEQ ID NOS: 14342  
; SEQ ID NO 5281  
; LENGTH: 210  
; TYPE: DNA  
; ORGANISM: Klebsiella pneumoniae  
US-09-489-039A-5281

Query Match 12.4%; Score 134; DB 4; Length 210;  
Best Local Similarity 80.5%; Pred. No. 1.2e-21;  
Matches 169; Conservative 0; Mismatches 40; Indels 1; Gaps 1;  
Qy 875 TCATGGGTGGGTGTTTCGTCATGGAACCTCTCGGTGATGATCCAGTCTGCTTCA 934  
Db 1 TCATGGCGGGGTTTTTCGTGGTGAACCGCTGCTGCTTATCTGCGGCTCCTTTA 60  
Qy 935 AGCTG-ACCGAGCGCGCTCTTCGTTATGGCGCGGATCCATCACCATTTCGACTGAA 993  
Db 61 AGCTGCGCGGCTGAGCGCATCTTCGCGATGGCGCGGATTCACACCATATGACTGAA 120  
Qy 994 GGTGTCGCGGACCGCGCGCTGATGCTGCTGCTTCTGATCATCACCCTGATCTGCTGCTG 1053  
Db 121 GGTGTCGCGGAGCGCGCTTATGTCGCTTCTGATTTTCGCTGATGCTGCTGCTG 180  
Qy 1054 ATCGGCTCTGCCACCTTGAGCTGCTTGA 1083

Db 181 ATTGGCCTGGGCGCTGAAGGTACGTAA 210  
RESULT 14  
US-08-311-174-4  
; Sequence 4, Application US/08311174  
; Patent No. 5556776  
; GENERAL INFORMATION:  
; APPLICANT: TSUCHIYA, MAKOTO  
; APPLICANT: MIWA, KIYOSHI  
; TITLE OF INVENTION: SUCRASE GENE DERIVED FROM CORYNEFORM  
; TITLE OF INVENTION: BACTERIA  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,  
; ADDRESSEE: P.C. Jefferson Davis Highway, Suite 400  
; CITY: Arlington  
; STATE: Virginia  
; COUNTRY: U.S.A.  
; ZIP: 22202  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/311,174  
; FILING DATE: 23-SEP-1994  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 046836/1992  
; FILING DATE: 04-MAR-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Oblon, No. 5556776man F.  
; REGISTRATION NUMBER: 24,618  
; REFERENCE/DOCKET NUMBER: 10-699-0  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703) 413-3000  
; TELEFAX: (703) 413-2220  
; TELEX: 248855 OPAT UR  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 6911 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: unknown  
; TOPOLOGY: unknown  
; MOLECULE TYPE: DNA (genomic)  
US-08-311-174-4

Query Match 12.3%; Score 133.4; DB 1; Length 6911;  
Best Local Similarity 54.9%; Pred. No. 3.7e-21;  
Matches 308; Conservative 0; Mismatches 246; Indels 7; Gaps 2;  
Qy 519 CATCTCTCTGCTGCTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 578  
Db 6079 CATCTGTTCTCTCATTTTATCTACCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6138  
Qy 579 CACCGACGCTCTGACGCGCTGCG-ATCATGCGGACGCTAATGTTGCTGCGCGC 632  
Db 6139 CATGATGTTTGGATGGTTTGGCTCCAGTACCACAGATTGTCATGCTGCTTACAC 6198  
Qy 633 GCTGGGCGATCTTCTGCTACCTGCTGCGGCAACGTAAGTTTCCGCGAGTACCTGCTGATTC 692  
Db 6199 CTTGATCACGCTCTGCGAGTTCCGAACTCTCCGATCTGCTGAGTGAAGCGGTTGCAA 6258  
Qy 693 CAAGTACCGGCGCGCGGAGCTGATCGTTTCTGCGCGGCTGCTGCTGCGCGCGCT 752  
Db 6259 TACGTCGCTGATCCACTGATTTGCTGTTTGTGCTGCTGCTGCTGCTGCTGCTGCT 6317  
Qy 753 CGGCTTCTGCTTCAACACCTATCCGCGCAGCTTTCATGGGCGAGCTTCGCGCGCT 812



GenCore version 5.1.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 9, 2004, 12:40:19 ; Search time 492 Seconds  
(without alignments)  
9351.217 Million cell updates/sec

Title: US-10-089-787-1  
Perfect score: 1083  
Sequence: 1 atgctctgctgctgcccga.....ccacctgaagctgcgttga 1083

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_29Jan04:\*

- 1: Geneseqn1980s:\*
- 2: Geneseqn1990s:\*
- 3: Geneseqn2000s:\*
- 4: Geneseqn2001as:\*
- 5: Geneseqn2001bs:\*
- 6: Geneseqn2002s:\*
- 7: Geneseqn2003as:\*
- 8: Geneseqn2003bs:\*
- 9: Geneseqn2003cs:\*
- 10: Geneseqn2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1083	100.0	1083	4	AAF86639 Pseudomon
2	1078.2	99.6	1083	4	AAS54276 Pseudomon
3	1078.2	99.6	1083	7	ACA42599 Prokaryot
4	819.8	75.7	1080	7	ACA43847 Prokaryot
5	767	70.8	1080	7	ACA45659 Prokaryot
6	540.8	49.9	1080	7	ACA31787 Prokaryot
7	539	49.8	1083	4	AAS53957 Klebsiell
8	537.6	49.6	1080	7	ACA35532 Prokaryot
9	518.2	47.8	1083	7	ACA51132 Prokaryot
10	518.2	47.8	1083	7	ACA51964 Prokaryot
11	517.2	47.8	1083	4	AAS56340 Salmonell
12	502.2	46.4	1083	4	AAS52294 E. coli D
13	502.2	46.4	1083	5	AAS81459 Escherich
14	502.2	46.4	1083	7	ACA32358 Prokaryot
15	490.6	45.3	1060	7	ACA49179 Prokaryot
16	480.4	44.4	1167	7	ACA27323 Prokaryot
17	470.2	43.4	1083	7	ACA53405 Prokaryot
18	460.6	42.5	1083	7	ACA53866 Prokaryot
19	440.4	40.7	1167	7	ACA25959 Prokaryot
20	412.6	38.1	1167	7	ACA25051 Prokaryot
21	410	37.9	1083	7	ACA42010 Prokaryot
22	410	37.9	1131	3	AAS53168 Neisseria
23	409.2	37.8	1080	7	AB241538 N. gonorr

24	409.2	37.8	1080	7	ACA41076 Prokaryot
25	406.2	37.5	1083	7	ACF71681 Photorhab
26	406.2	37.5	110000	7	Continuation (49 o
27	406.2	37.5	110000	7	ACF6387 Photorhab
28	405.2	37.4	69936	3	AAAB1479 N. mening
29	405.2	37.4	110000	3	Continuation (5 of
30	405.2	37.4	349980	3	AAF21607 Neisseria
31	402	37.1	1131	3	AAS53166 Neisseria
32	388.6	35.9	1083	7	ACA42842 Prokaryot
33	368	34.0	1080	7	ACA44760 Prokaryot
34	361.4	33.4	1083	4	AAS53452 Haemophil
35	361.4	33.4	1083	7	ACA4365 Prokaryot
36	361.4	33.4	110000	2	ACA4365 Prokaryot
37	361.4	33.4	110000	2	AA742063_11
38	359	33.1	4942	5	AAS89816 DNA encod
39	359	33.1	4942	5	AAS89273 DNA encod
40	359	33.1	4944	5	AAS90181 DNA encod
41	326.6	30.2	1310	5	AAS93847 DNA encod
42	326	30.1	645	7	ACA24034 Prokaryot
43	310	28.6	1257	5	AAS90011 DNA encod
44	294.6	27.2	433	7	AB256666 Aspergill
45	279.8	25.8	1131	3	AAS53167 Neisseria

## ALIGNMENTS

RESULT 1  
AAF86639  
ID AAF86639 standard; DNA; 1083 BP.  
XX  
AC AAF86639;  
XX  
DT 17-JUL-2001 (first entry)  
XX  
DE Pseudomonas aeruginosa mray gene.  
XX  
KW Mray, phospho-N-acetylmuramoyl-pentapeptide translocase;  
KW peptidoglycan biosynthesis; bacterial cell wall; infection;  
KW drug screening; antibacterial; Pseudomonad; G+C rich bacterium; ds.  
XX  
OS Pseudomonas aeruginosa.  
XX  
FH Key Location/Qualifiers  
FT CDS 1..1083  
FT /tag= a  
FT /product= "Mray protein"  
FT /function= "Catalyses the transfer of N-acetylmuramic  
FT acid peptide to a bactoprenol phosphate carrier molecule  
FT in peptidoglycan biosynthesis"

WO200125251-A1.

12-APR-2001.

29-SEP-2000; 2000WO-US027056.

04-OCT-1999; 99US-0157580P.

(MERI ) MERCK & CO INC.

El-Sherbeini M, Azzolina B;

WPI; 2001-308221/32.

P-PSDB; AAB73487.

New Mray gene and enzyme of Pseudomonas aeruginosa, useful in vitro assays for screening antibacterial compounds that target cell wall biosynthesis, particularly for screening antibiotics against Pseudomonads.

Claim 2; Fig 1; 22pp; English.

This sequence represents the Pseudomonas aeruginosa mray gene, which encodes a phospho-N-acetylmuramyl-pentapeptide translocase designated Mray. The Mray protein is involved in bacterial cell wall biosynthesis. It catalyzes the first step of the membrane cycle of peptidoglycan biosynthesis, i.e., the transfer of an N-acetylmuramic acid peptide to a bactoprenol phosphate carrier molecule. The Mray protein is useful in vitro assays to screen for antibacterial compounds that target cell wall biosynthesis. Inhibitors of the Mray protein are useful in preventing the growth of Pseudomonas and other G+C rich bacteria. Pseudomonas aeruginosa is an opportunistic pathogen which causes infections in patients with burns, neutropenia, or cystic fibrosis. Primers or probes derived from the mray gene are useful in nucleic acid amplification-based assays for detecting the presence of a polynucleotide encoding Pseudomonas aeruginosa Mray protein

Sequence 1083 BP; 164 A; 350 C; 333 G; 236 T; 0 U; 0 Other;

Query Match 100.0%; Score 1083; DB 4; Length 1083;  
Best Local Similarity 100.0%; Pred. No. 3.5e-209;  
Matches 1083; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGCTCCTGCTGCTGCGGATACCTGCAACAGTTCTTCAAGGGCTTCGGGCTTTCCAG 60  
DB 1 ATGCTCCTGCTGCTGCGGATACCTGCAACAGTTCTTCAAGGGCTTCGGGCTTTCCAG 60  
QY 61 TACCTGACCTCGCGGATTTCTACGCGTGTCTACGCGTGTCTGCTGCTGCTGCTG 120  
DB 61 TACCTGACCTCGCGGATTTCTACGCGTGTCTACGCGTGTCTGCTGCTGCTGCTG 120  
QY 121 GGGCCCTGATGATCGTACCTTTGAGATCCCGGATCGGCGGCTGCGCAACGAC 180  
DB 121 GGGCCCTGATGATCGTACCTTTGAGATCCCGGATCGGCGGCTGCGCAACGAC 180  
QY 181 GGTGCGCAGTCTGCTGCAAGAGGACCGGACCGGATCGGCGGCTGATCTT 240  
DB 181 GGTGCGCAGTCTGCTGCAAGAGGACCGGACCGGATCGGCGGCTGATCTT 240  
QY 241 ACGCCATAGCATCAGACGCTGCTGCGGCGATCTTCCAAACGCTACGTGGGTA 300  
DB 241 ACGCCATAGCATCAGACGCTGCTGCGGCGATCTTCCAAACGCTACGTGGGTA 300  
QY 301 GTGCTGGTACCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360  
DB 301 GTGCTGGTACCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360  
QY 361 ATCGAAGACCTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420  
DB 361 ATCGAAGACCTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420  
QY 421 GGCATCGGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480  
DB 421 GGCATCGGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480  
QY 481 ATCGTCCGATGCTGAGAGCGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCC 540  
DB 481 ATCGTCCGATGCTGAGAGCGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCC 540  
QY 541 TACTTGTCTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600  
DB 541 TACTTGTCTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600  
QY 601 GCGATCATGCGGAGGTAATGGTTCGCGGCGCTGGGATCTTCTGCTACCTGCTGGG 660  
DB 601 GCGATCATGCGGAGGTAATGGTTCGCGGCGCTGGGATCTTCTGCTACCTGCTGGG 660  
QY 661 AACGTGAAGTTCGCGGAGTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720  
DB 661 AACGTGAAGTTCGCGGAGTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720  
QY 721 GTGTTTCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780  
DB 721 GTGTTTCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780

QY 781 GGCACGGTCTTCATGGCGAGTGGCGGCTGGCGGCTGGCGGCTGGCGGCTGGCGGCT 840  
DB 781 GGCACGGTCTTCATGGCGAGTGGCGGCTGGCGGCTGGCGGCTGGCGGCTGGCGGCT 840  
QY 841 GGGGTGATCGTGGCGGAGAGATCGTGTGTTTCAATGCTGGTGGGTTTCTCATGGAA 900  
DB 841 GGGGTGATCGTGGCGGAGAGATCGTGTGTTTCAATGCTGGTGGGTTTCTCATGGAA 900  
QY 901 ACCCTCTCGTATGATCAGTCCGCTTCTTCAAGCTGACCGGACGGCGGCTTCCGT 960  
DB 901 ACCCTCTCGTATGATCAGTCCGCTTCTTCAAGCTGACCGGACGGCGGCTTCCGT 960  
QY 961 ATGGGCGCGATCCATCACCATTTCGAAGCTGCGCGGACCGCGGCTGATCGTG 1020  
DB 961 ATGGGCGCGATCCATCACCATTTCGAAGCTGCGCGGACCGCGGCTGATCGTG 1020  
QY 1021 CGCTTCTGATCATCACCGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1080  
DB 1021 CGCTTCTGATCATCACCGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1080  
QY 1081 TGA 1083  
DB 1081 TGA 1083

RESULT 2  
AAS54276

ID AAS54276 standard; DNA; 1083 BP.

XX AAS54276;

DT 13-FEB-2002 (first entry)

DE Pseudomonas aeruginosa DNA for cellular proliferation protein #407.

XX Antisense; ds; prokaryotic cellular proliferation gene; antibiotic;  
KW antibacterial; drug design.

XX Pseudomonas aeruginosa.

XX WO200170955-A2.

PD 27-SEP-2001.

XX 21-MAR-2001; 2001WO-US009180.

PR 21-MAR-2000; 2000US-0191078P.

PR 23-MAY-2000; 2000US-0206848P.

PR 26-MAY-2000; 2000US-0207727P.

PR 23-OCT-2000; 2000US-0242578P.

PR 27-NOV-2000; 2000US-0253625P.

PR 22-DEC-2000; 2000US-0257931P.

PR 16-FEB-2001; 2001US-0269308P.

XX (ELIT-) ELITRA PHARM INC.

XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;  
PI Yamamoto RT, Xu HH;

XX WPI; 2001-611495/70.

XX P-PSDB; AAU36417.

XX New polynucleotides for the identification and development of  
PT antibiotics, comprise sequences of antisense nucleic acids.

XX Claim 27; SEQ ID NO 7913; Slipp; English.

XX The invention relates to antisense inhibitors of genes essential to  
CC prokaryotic cellular proliferation, their use in identifying the genes,  
CC their use in the discovery of novel antibiotics, the essential genes  
CC themselves and the encoded proteins. The prokaryotes used are Escherichia  
CC coli, Staphylococcus aureus, Salmonella typhi, Klebsiella pneumoniae,  
CC Pseudomonas aeruginosa and Enterococcus faecalis. The invention is also

useful for the identification of potential new targets for antibiotic development. The antisense nucleic acids can also be used to identify proteins used in proliferation, to express these proteins, and to obtain antibodies capable of binding to the expressed proteins. The proteins can be used to screen compounds in rational drug discovery programs. The antisense nucleic acid sequence is also useful to screen for homologous nucleic acids which are required for cell proliferation in a wide variety of organisms. The present sequence encodes an essential prokaryotic cellular proliferation protein. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

Sequence 1083 BP; 163 A; 348 C; 336 G; 236 T; 0 U; 0 Other;

Query Match 99.6%; Score 1078.2; DB 4; Length 1083;  
 Best Local Similarity 99.7%; Pred. No. 3.3e-208;  
 Matches 1080; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ATGCTCTGCTGTCGCGGATACCTGCAACAGTCTACAAGGGCTTCGGGCTTCCTCAG 60  
 DB 1 ATGCTCTGCTGTCGCGGATACCTGCAACAGTCTACAAGGGCTTCGGGCTTCCTCAG 60

QY 61 TACCTGACCTGCGCGGCAATCTCAGCGTCTCAGCGGCTGCGGTGCTGCGTGGCTG 120  
 DB 61 TACCTGACCTGCGCGGCAATCTCAGCGTCTCAGCGGCTGCGGTGCTGCGTGGCTG 120

QY 121 GGGCCCTGATGATCGGTACCTTGCAGATCCCGAGTCCCGAGCGGCTGCGCAACGAC 180  
 DB 121 GGGCCCTGATGATCGGTACCTTGCAGATCCCGAGTCCCGAGCGGCTGCGCAACGAC 180

QY 181 GGTCCGAGTCGCACTGTGCGAAGAGGGCACCCGACCATGCGCGCGCGCTGATCCTT 240  
 DB 181 GGTCCGAGTCGCACTGTGCGAAGAGGGCACCCGACCATGCGCGCGCGCTGATCCTT 240

QY 241 ACCGCCATAGCCATGACAGCGTCTGTGGGCGGATCTTCCAAACGCTACGTTGGGTA 300  
 DB 241 ACCGCCATAGCCATGACAGCGTCTGTGGGCGGATCTTCCAAACGCTACGTTGGGTA 300

QY 301 GTGCTGTGCTTACCTTCTGCTGCGTCCATCGGCTGGGTAGACCATCCGCAAGGTG 360  
 DB 301 GTGCTGTGCTTACCTTCTGCTGCGTCCATCGGCTGGGTAGACCATCCGCAAGGTG 360

QY 361 ATCGAGAAGAACTCCGCTGCGGCTGCGAGCGTCTGGAAGTACTTCTGGCAGTCTGGTTC 420  
 DB 361 ATCGAGAAGAACTCCGCTGCGGCTGCGAGCGTCTGGAAGTACTTCTGGCAGTCTGGTTC 420

QY 421 GGCATCGCGCGCGGCTGCTTCTCTACATGCTGCGAAGACCCGATCGAGACACCGTG 480  
 DB 421 GGCATCGCGCGCGGCTGCTTCTCTACATGCTGCGAAGACCCGATCGAGACACCGTG 480

QY 481 ATCGTCCGATGCTGAAGAGCGTTCGAGATCCAGTTGGGCATCTTCTGCTGCTGAC 540  
 DB 481 ATCGTCCGATGCTGAAGAGCGTTCGAGATCCAGTTGGGCATCTTCTGCTGCTGAC 540

QY 541 TACTTGTCTATCGTCCGCTGCGAATGCGGTGACCTCAGCAGCGTCTCGAGCGCTG 600  
 DB 541 TACTTGTCTATCGTCCGCTGCGAATGCGGTGACCTCAGCAGCGTCTCGAGCGCTG 600

QY 601 GCGATCATGCGCGGTAATGTTGTCGCGCGCTGGGCATCTTCTGCTACCTGTGGGC 660  
 DB 601 GCGATCATGCGCGGTAATGTTGTCGCGCGCTGGGCATCTTCTGCTACCTGTGGGC 660

QY 661 AACGTGAAGTTCGCGGATGCTGCTGATTTCCCAACGTACGGGCGCGCGAGCTGATC 720  
 DB 661 AACGTGAAGTTCGCGGATGCTGCTGATTTCCCAACGTACGGGCGCGCGAGCTGATC 720

QY 721 GTGTTCTGCGCGCGGCTGCTGCGCGCGCTCGGCTTCTGCTGTTCAACACCTATCCG 780  
 DB 721 GTGTTCTGCGCGCGGCTGCTGCGCGCGCTCGGCTTCTGCTGTTCAACACCTATCCG 780

QY 781 GCGGAGTCTTCAAGGGCGAGTGGCGCGTGGCGTGGGCGCGGCTGGGCAACATC 840  
 DB 781 GCGGAGTCTTCAAGGGCGAGTGGCGCGTGGCGTGGGCGCGGCTGGGCAACATC 840

781 GCGCAGGTCTTCATGGCGGAGCTGCGGCGGCTGGCGGCTGGCGGCGGCTGGGACCATC 840  
 841 GCGGTGATCGTCCGCCAGGAGATCGTCTGTTTCATCATGGGTGGGTGTTGTTGATGAA 900  
 841 GCGGTGATCGTCCGCCAGGAGATCGTCTGTTTCATCATGGGTGGGTGTTGTTGATGAA 900  
 901 ACCCTCGGTGATGATCCAGGCTGCTTCTTCAAGCTGACCGGACCGCGGCTTTCCT 960  
 901 ACCCTCGGTGATGATCCAGGCTGCTTCTTCAAGCTGACCGGACCGCGGCTTTCCT 960  
 961 ATGGCGCGGATCCATCACCATTTCGAACTGAAAGCTGGCGGACCGCGGCTGATCGT 1020  
 961 ATGGCGCGGATCCATCACCATTTCGAACTGAAAGCTGGCGGACCGCGGCTGATCGT 1020  
 1021 GCTTCTGGATCATCAGGTGATCTGCTGCTGATCGGCTTCGACCATTTGAAGCTGCT 1080  
 1021 GCTTCTGGATCATCAGGTGATCTGCTGCTGATCGGCTTCGACCATTTGAAGCTGCT 1080

1081 TGA 1083  
 1081 TGA 1083

RESULT 3  
 ACA42599  
 ID ACA42599 standard; DNA; 1083 BP.

XX ACA42599;  
 XX AC  
 XX DT 19-JUN-2003 (first entry)  
 XX DE Prokaryotic essential gene #24256.  
 XX KW Antisense; ds; prokaryotic essential gene; cell proliferation;  
 XX KW drug design; gene.  
 XX OS Pseudomonas aeruginosa.  
 XX PN WO200277183-A2.  
 XX PD 03-OCT-2002.  
 XX PF 21-MAR-2002; 2002WO-US009107.  
 XX PR 21-MAR-2001; 2001US-00815242.  
 XX PR 06-SEP-2001; 2001US-00948993.  
 XX PR 25-OCT-2001; 2001US-0342923P.  
 XX PR 08-FEB-2002; 2002US-00072851.  
 XX PR 06-MAR-2002; 2002US-0362699P.  
 XX PA (ELIT-) ELITRA PHARM INC.  
 XX PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;  
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;  
 XX WPI: 2003-029926/02.  
 XX P-PSDB; ABU38729.  
 XX PT New antisense nucleic acids, useful for identifying proteins or screening  
 XX PT for homologous nucleic acids required for cellular proliferation to  
 XX PT isolate candidate molecules for rational drug discovery programs.  
 XX PS Claim 14; SEQ ID NO 30469; 1766pp; English.  
 XX CC The invention relates to an isolated nucleic acid comprising any one of  
 CC the 6213 antisense sequences given in the specification where expression  
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:  
 CC (1) a vector comprising a promoter operably linked to the nucleic acid  
 CC encoding a polypeptide whose expression is inhibited by the antisense  
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated  
 CC polypeptide or its fragment whose expression is inhibited by the  
 CC antisense nucleic acid; (4) an antibody capable of specifically binding  
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular



XX The invention relates to an isolated nucleic acid comprising any one of  
CC the 6213 antisense sequences given in the specification where expression  
CC of the nucleic acid inhibits proliferation of a cell. Also included are:  
CC (1) a vector comprising a promoter operably linked to the nucleic acid  
CC encoding a polypeptide whose expression is inhibited by the antisense  
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated  
CC polypeptide or its fragment whose expression is inhibited by the  
CC antisense nucleic acid; (4) an antibody capable of specifically binding  
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular  
CC proliferation or the activity of a gene in an operon required for  
CC the gene product or that has an activity against a biological pathway  
CC required for proliferation, or that inhibits cellular proliferation; (8)  
CC identifying a gene required for cellular proliferation or the biological  
CC pathway in which a proliferation-required gene or its gene product lies  
CC on a gene on which the test compound that inhibits proliferation of an  
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a  
CC compound's activity; (11) a culture comprising strains in which the gene  
CC product is overexpressed or underexpressed; (12) determining the extent  
CC to which each of the strains is present in a culture or collection of  
CC strains; or (13) identifying the target of a compound that inhibits the  
CC proliferation of an organism. The antisense nucleic acids are useful for  
CC identifying proteins or screening for homologous nucleic acids required  
CC for cellular proliferation to isolate candidate molecules for rational  
CC drug discovery programs, or for screening homologous nucleic acids  
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,  
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target  
CC prokaryotic essential genes. Note: The sequence data for this patent did  
CC not form part of the printed specification, but was obtained in  
CC electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 1080 BP; 169 A; 329 G; 334 G; 248 T; 0 U; 0 Other;  
  
Query Match 75.7%; Score 819.8; DB 7; Length 1080;  
Best Local Similarity 85.0%; Pred. No. 4.3e-156;  
Matches 917; Conservative 0; Mismatches 162; Indels 0; Gaps 0;  
  
QY 1 ATGCTCTGCTGCTGGCGGAACTCCTGCAACAGTCTTCAAGGCTTGGCGTCTCCAG 60  
DB 1 ATGCTGCTGCTGCTGGCGGAACTCCTGCAACAGTCTTCAAGGCTTGGCGTCTCCAG 60  
  
QY 61 TACTGACCTGGCGGCACTTCTCAGCGTCTCACCGCGCTGTCGCTGCTGCTGCTG 120  
DB 61 TACTGTCCTGGCGGCACTTCTGCTGCTACGCGCTGTCGCTGCTGCTGCTGCTG 120  
  
QY 121 GGGCCCTGGATGATCCGTTACCTGATCCCGAGATCCCGAGATCCCGAGATCCCGAG 180  
DB 121 GGGCCCTGGATGATCCGTTACCTGATCCCGAGATCCCGAGATCCCGAGATCCCGAG 180  
  
QY 181 GGTCCGAGTCGACCTGTGCAAGAGAGGCAACCCGACCATGGCGGCGCTGATCCTT 240  
DB 181 GGTCCGCAATCGACCTGTGCAAGTCCGCGCAACCCGACCATGGTGGGCGCTGATCCTG 240  
  
QY 241 ACCGCGATAGCATCAGCAGCTGCTGTGGGCGGATCTTCCAAACGCTACGTTGGGTA 300  
DB 241 TCGGCCATCGCGCTCAGCAGCTTCTGCTGGGCGGACCTGAGCAACCGCTATGATGGGTT 300  
  
QY 301 GTGCTGCTGCTTACCTGCTGTTCCGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360  
DB 301 GTGCTGATGCTACCTTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360  
  
QY 361 ATCAGAGAAGAACTCCCGTGGCTGCCAGCGCTGGAAGTACTTCTGGCAGTGGGTTC 420  
DB 361 ATCGAAAGAAGAACTCCCGTGGCTGCCAGCGCTGGAAGTACTTCTGGCAATCGGTTC 420  
  
QY 421 GGCAATCGCGCGCGGCTGCTTCTACATGCTGCGGAAACCGCATCCGAGACCCCTG 480  
DB 421 GGCTTCGGCGCGGCTGCTTCTGCTACAGACGCGCAACAGCGTCCGAGACCGCTG 480  
  
QY 481 ATCGTCCGATGCTGAAGAGCGTGCAGATCCAGTTGGGCACTCTTCTGCTGCTGCTG 540  
DB 481 ATCGTCCGATGCTGAAGAGCGTGCAGATCCAGTTGGGCACTCTTCTGCTGCTGCTG 540

DB 481 ATCGTCCGCTTCATCAAGGATGTCAACATTCGCTGGCGCTGCTGCTGCTGCTGCTG 540  
QY 541 TACTTCTGCTACCTGCTGGCTCGAGCAATGCGGTGAACTCACCAGCGGTCTCGAGCGGCTG 600  
DB 541 TACTTCTGCTACCTGCTGGCTCGAGCAATGCGGTGAACTCACCAGCGGTCTCGAGCGGCTG 600  
QY 601 CGGATCATGCCAGAGTAAATGTTGCGGCGCGCTGGGCACTCTTCTGCTACCTGCTGCGGC 660  
DB 601 GCAATCATGCCAGAGTAAATGTTGCGGCGCGCTGGGCACTCTTCTGCTACCTGCTGCGGT 660  
QY 661 AACGTGAAGTTTCGCGGAGTACTGCTGATTCCTCCAAAGTACCGGCGCGGCGAGTATC 720  
DB 661 AACGTGAAGTTTCGCGGAGTACTGCTGATTCCTCCAAAGTACCGGCGCGGCGAGTATC 720  
QY 721 GTGTTCTGCGCGCGCTGGTTCGCGCGCGCTCGGCTTCTGCTGCTGCTGCTGCTGCTGCTG 780  
DB 721 GTGTTCTGCGCGCGCTGGTTCGCGCGCGCTCGGCTTCTGCTGCTGCTGCTGCTGCTGCTG 780  
QY 781 GCGCAGGCTTTCATGGGCGAGTCTGCGCGCGCTGCGCTGCGCGCGCTGCGCGCGCTGCG 840  
DB 781 GCGCAGGCTTTCATGGGCGAGTCTGCGCGCGCTGCGCTGCGCGCGCTGCGCGCGCTGCG 840  
QY 841 GCGGTGATGTCGCGCGAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900  
DB 841 GCGGTGATGTCGCGCGAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900  
QY 901 ACCCTCTGCTGATGATCCAGTCTGCTTCTTCAAGTACCGGCGCGCTGCTGCTGCTGCTG 960  
DB 901 ACCCTCTGCTGATGATCCAGTCTGCTTCTTCAAGTACCGGCGCGCTGCTGCTGCTGCTG 960  
QY 961 ATGGCGCGGATTCATCACCCTTTGAACTGAAAGGCTGGCGCGCGCGCGCGCTGATCGTG 1020  
DB 961 ATGGCGCGGATTCATCACCCTTTGAACTGAAAGGCTGGCGCGCGCGCGCGCTGATCGTG 1020  
QY 1021 CGCTTCTGATGATCAGCGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1079  
DB 1021 CGCTTCTGATGATCAGCGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1079  
  
RESULT 5  
AC45659 standard; DNA; 1080 BP.  
XX AC AC45659;  
XX AC AC45659;  
DT 19-JUN-2003 (first entry)  
XX  
DE Prokaryotic essential gene #27316.  
XX  
KW Antisense; ds; prokaryotic essential gene; cell proliferation;  
KW drug design; gene.  
OS Pseudomonas syringae.  
XX  
XX  
PN WO200277183-A2.  
XX  
PD 03-OCT-2002.  
XX  
PF 21-MAR-2002; 2002WO-US009107.  
XX  
PR 21-MAR-2001; 2001US-00815242.  
PR 06-SEP-2001; 2001US-00948993.  
PR 25-OCT-2001; 2001US-0342923F.  
PR 08-FEB-2002; 2002US-00072851.  
PR 06-NAR-2002; 2002US-0362699P.  
XX  
PA (ELIT-) ELITRA PHARM INC.  
XX  
XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlson KU, Zyskind JW;  
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;  
XX WPI; 2003-029926/02.  
DR P-PSDB; ABU41789.

XX New antisense nucleic acids, useful for identifying proteins or screening  
PT for homologous nucleic acids required for cellular proliferation to  
PT isolate candidate molecules for rational drug discovery programs.  
XX  
XX  
PS Claim 14; SEQ ID NO 33529; 1766pp; English.  
XX

CC The invention relates to an isolated nucleic acid comprising any one of  
CC the 6213 antisense sequences given in the specification where expression  
CC of the nucleic acid inhibits proliferation of a cell. Also included are:  
CC (1) a vector comprising a promoter operably linked to the nucleic acid  
CC encoding a polypeptide whose expression is inhibited by the antisense  
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated  
CC polypeptide or its fragment whose expression is inhibited by the  
CC antisense nucleic acid; (4) an antibody capable of specifically binding  
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular  
CC proliferation or the activity of a gene in an operon required for  
CC proliferation; (7) identifying a compound that influences the activity of  
CC the gene product or that has an activity against a biological pathway  
CC required for proliferation, or that inhibits cellular proliferation; (8)  
CC identifying a gene required for cellular proliferation or the biological  
CC pathway in which a proliferation-required gene or its gene product lies  
CC or a gene on which the test compound that inhibits proliferation of an  
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a  
CC product is overexpressed or underexpressed; (12) determining the extent  
CC to which each of the strains is present in a culture or collection of  
CC strains; or (13) identifying the target of a compound that inhibits the  
CC proliferation of an organism. The antisense nucleic acids are useful for  
CC identifying proteins or screening for homologous nucleic acids required  
CC for cellular proliferation to isolate candidate molecules for rational  
CC drug discovery programs, or for screening homologous nucleic acids  
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,  
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target  
CC prokaryotic essential genes. Note: The sequence data for this patent did  
CC not form part of the printed specification, but was obtained in  
CC electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences

XX Sequence 1080 BP; 168 A; 302 C; 330 G; 280 T; 0 U; 0 Other;

Query Match 70.8%; Score 767; DB 7; Length 1080;  
Best Local Similarity 81.9%; Pred. No. 1.9e-145;  
Matches 884; Conservative 0; Mismatches 195; Indels 0; Gaps 0;

QY 1 ATGCTCTGCTGCTGCGGCAATACCTGCAACAGTCTTACAAGGGCTTCGGGCTCTTCAG 60  
DB 1 ATGCTGCTGCTGCTGCGGCAATACCTGCAACAGTCTTACAAGGGCTTCGGGCTCTTCAG 60  
QY 61 TACCTGACCTGCGGGGATTCACGCTGCTCAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120  
DB 61 TACCTGCTGCTGCGGGGATTCCTGCGGTGCTCAGTGGCTGACGTTGCTGCTGCTGCTGCTG 120  
QY 121 GGGCCCTGGATGATCCGTACCTTGAGATGCCCGGAGATGCCCGGAGATGCCCGGAGATGCC 180  
DB 121 GGGCCCTGGATGATCCGACCTTGAGATGCCCGGAGATGCCCGGAGATGCCCGGAGATGCC 180  
QY 181 GGTCCGAGTCCGACCTGCTGAGAGAGGACCCGACCATGGCGGCGGCGCTGATCTT 240  
DB 181 GGTCCGAGTCCGACCTGCTGAGAGAGGACCCGACCATGGCGGCGGCGCTGATCTT 240  
QY 241 ACCGCCATAGCCATCAGACGCTGCTGCTGCGGGGATCTTCCAAACGCTGCTGCTGCTGCTG 300  
DB 241 TCGTCCATCGGTATCAGACCTTGTCTGCGGTGATTTGAGCAACCGTTACGTTGCTGCTG 300  
QY 301 GTGCTGGTGTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360  
DB 301 GTGCTGGTGTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360  
QY 361 ATCGAAGAAGTCCCGTGGCTGCGGAGCCGCTGGAAGTCTTCTGGCAGTCCGCTGCTGCTG 420  
DB 361 ATTGAAGAAGTCCCGTGGCTGCGGAGCCGCTGGAAGTCTTCTGGCAGTCCGCTGCTGCTG 420

QY 421 GGCAATCGGCGCGCGCTGCTTCTTACATGATCGCGGAAACCCCGATGAGACACCCCTG 480  
DB 421 GGCAATCGGCGCGCGCTTCTTCTGACGATCGCGGAAACCCCGATGAGACACCACTG 480  
QY 481 ATCGTCCGATGCTGAAGAGCGTCCAGATCCAGTTCAGTTCGGGATCTTCTGCTGCTGCTG 540  
DB 481 ATCGTCCGATGCTGAAGAGCGTCCAGATTCGGCTCGGATCGGCTTCATCGTCTGCTG 540  
QY 541 TACTTCGTCATCGTGGCTGCGAGCAATCGGCTGAGTCACTCAGGAGCTTCGACCGGCTG 600  
DB 541 TACTTCGTCATCGTGGCTTCCAGCAACGCACTCACTGAGTCACTGAGCGGCTGAGCGGCTG 600  
QY 601 GCGATCATCCGACGCTAATGCTGCGCGCGCTGCGGATCTTCTGCTGCTGCTGCTGCTG 660  
DB 601 GCGATCATCCGACGCTGATGCTGCGCGCTGCGGATCTTCTGCTGCTGCTGCTGCTG 660  
QY 661 AAGTGAATTCGCGAGTACTGCTGATTCCTCAACGATACCGGCGCGGCTGAGCTGATC 720  
DB 661 AAGTGAATTCGCGGATACCTGCTGATTCCTCAACGATACCGGCGCGGCTGAGCTGATC 720  
QY 721 GTGTTCTGCGCGCTGCTGCTGCGCGCGCTGCGGATCTTCTGCTGCTGCTGCTGCTG 780  
DB 721 GTGTTCTGCGGCTGCTGATGCTGCGCGCGCTGCGGATCTTCTGCTGCTGCTGCTGCTG 780  
QY 781 GCGAGGCTCTTCAATGCGCGGATGCTGCGCGCGCTGCGGATCTTCTGCTGCTGCTGCTG 840  
DB 781 GCTCAGGCTGTTCAATGCGCGGATGCTGCGCGCGCTGCGGATCTTCTGCTGCTGCTGCTG 840  
QY 841 GCGGTCATGCTGCGCGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900  
DB 841 GCGGTCATGCTGCGGCTGAGGAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900  
QY 901 ACCCTCTGCTGATGATCCAGTCTGCTTCAAGCTGACCGGAGCGGCTGCTGCTGCTGCTG 960  
DB 901 ACCCTTTCAGTCTGATTCAGGCTGCTTCAAGCTGACCGGAGCGGCTGCTGCTGCTGCTG 960  
QY 961 ATGCGCGGATTCATCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020  
DB 961 ATGCGCGGATTCATCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020  
QY 1021 CGCTTCTGATCATCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1079  
DB 1021 CGCTTCTGATCATCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1079  
RESULT 6  
ACA31787  
ID ACA31787 standard; DNA; 1080 BP.  
XX ACA31787;  
XX  
XX 19-JUN-2003 (first entry)  
XX  
XX Prokaryotic essential gene #13444.  
XX Antisense; ds; prokaryotic essential gene; cell proliferation;  
XX drug design; gene.  
XX Enterobacter cloacae.  
XX OS  
XX PN W0200277183-A2.  
XX  
XX 03-OCT-2002.  
XX  
XX 21-MAR-2002; 2002WO-US0009107.  
XX  
XX 21-MAR-2001; 2001US-00815242.  
XX 06-SEP-2001; 2001US-00948993.  
XX 25-OCT-2001; 2001US-0342923P.  
XX 08-FEB-2002; 2002US-00072851.  
XX 06-MAR-2002; 2002US-0362699P.  
XX (ELIT-) ELITRA PHARM INC.  
PA

XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen K, Zyskind JW;  
PI Wall D, Trawick JD, Carr GU, Yamamoto R, Forsyth RA, Xu HH;  
XX WPI; 2003-029926/02.  
DR P-PSDB; ABU27917.  
XX New antisense nucleic acids, useful for identifying proteins or screening  
PT for homologous nucleic acids, required for cellular proliferation to  
PT isolate candidate molecules for rational drug discovery programs.  
XX Claim 14; SEQ ID NO 19657; 1766pp; English.  
XX The invention relates to an isolated nucleic acid comprising any one of  
CC the 6213 antisense sequences given in the specification where expression  
CC of the nucleic acid inhibits proliferation of a cell. Also included are:  
CC (1) a vector comprising a promoter operably linked to the nucleic acid  
CC encoding a polypeptide whose expression is inhibited by the antisense  
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated  
CC antisense nucleic acid; (4) an antibody capable of specifically binding  
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular  
CC proliferation or the activity of a gene in an operon required for  
CC proliferation; (7) identifying a compound that influences the activity of  
CC the gene product or that has an activity against a biological pathway  
CC required for proliferation, or that inhibits cellular proliferation; (8)  
CC identifying a gene required for cellular proliferation or the biological  
CC pathway in which a proliferation-required gene or its gene product lies  
CC or a gene on which the test compound that inhibits proliferation of an  
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a  
CC compound's activity; (11) a culture comprising strains in which the gene  
CC product is overexpressed or underexpressed; (12) determining the extent  
CC to which each of the strains is present in a culture or collection of  
CC strains; or (13) identifying the target of a compound that inhibits the  
CC proliferation of an organism. The antisense nucleic acids are useful for  
CC identifying proteins or screening for homologous nucleic acids required  
CC for cellular proliferation to isolate candidate molecules for rational  
CC drug discovery programs, or for screening homologous nucleic acids  
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,  
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target  
CC prokaryotic essential genes. Note: The sequence data for this patent did  
CC not form part of the printed specification, but was obtained in  
CC electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 1080 BP; 169 A; 283 C; 334 G; 294 T; 0 U; 0 Other;  
Query Match 49.9%; Score 540.8; DB 7; Length 1080;  
Best Local Similarity 68.8%; Pred. No. 7.9e-100;  
Matches 743; Conservative 0; Mismatches 337; Indels 0; Gaps 0;  
QY 1 ATGCTCCTGCTGCTGGCCGAATACCTGCAACAGTTCTACAGGGCTTCGGCGTCTTCCAG 60  
DB 1 ATGTTAGTTTGGCTGGCCGAACATTTGGTCAATATATTATTCAGGCTTTAACGTTCTTTTCG 60  
QY 61 TACTGACCTGCGCGCATCTTCAGCGTGTCTCAGCGGTCTGCTGCTGCTGCTGCTGCTG 120  
DB 61 TATCTGACGTTTCGGCCATCGTCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120  
QY 121 GGSCCTCGATGATCCGTACCTTCAGATCCCCAGATCGCCAGCGCGCTGCGCAACGAC 180  
DB 121 GGCCCGCGCATGATGCGCGCTCTGCAAAACTCTCTTTGGCCAGGTTGTGGTAAACGAC 180  
QY 181 GGTCCGAGTCGACCTGTGCAAGAGCGCACCCGACATGAGCGCGCGCTGATCCCTT 240  
DB 181 GGTCCGGAATCGCACTTCAGCAAGCGCGCACCGACGATGGCGGATCATGATCCCTG 240  
QY 241 ACCGCCATAGCCATCAGCAGCTGCTGTGGCGGATCTTTCCAAACCGCTACGTTGGGTA 300  
DB 241 ACCGCCATCGTGTGTTTCGGTGTGCTGTGGCGGTATCCCTCCAAACCGCTGCTGTC 300  
QY 301 GTGCTGTGCTTACCTCTGTTGCGTGCCATCGCTGGGTAGACGACTACCCCAAGTG 360

DB 301 GTTCTGACGGTACTGGTGGGTTACGGGATCATCGGTTTGTGTTGATGATTACCGCAAGTC 360  
QY 361 ATCGAGAGAACTCCCGTGGCTCCCGAGCGCTGGAAGTACTTCTGGAGTCCGTTGTC 420  
DB 361 GTTCGCAAGATACATAAGGGCTGATCGCGGCTGGAATATTTCTGGATGCTCGGTGATC 420  
QY 421 GGCATCGCGCGCGGCTTCTCTCATGATCTCCGAAACCCCGATCGAGACCAACCTG 480  
DB 421 CGCTGGCGGTGGCTTCGCACTGTATCTGGCAGGCAAGACACGCCAGCCAGCTG 480  
QY 481 ATCTGCGGATGCTGAAGAGCGTCAAGTCAAGTGGGATCTTTCTTCTGGTGTCTGACC 540  
DB 481 GTCGTGCGGTTCTTTAAAGGACGTGATCGCGGCTGTTTCTATCATCTCTGCGCC 540  
QY 541 TACTTCTGTCGCGCTCGAGCAATGCGGTGAACCTCACCGACGCTCTCGAGCGCTG 600  
DB 541 TACTTGTGATGTTGGTACCGGCAACCGGTAACTGACCGACGCTGATGATGCTG 600  
QY 601 CGCATCATCGCGACGTAATGTTGCGCGGCGCTGCGCATCTTCTCTACCTGTTCGGC 660  
DB 601 CGCATTAATGCCAACCGTCTTTGTGCGCGCGGTTTTCGCTGGTGGCGACCGCT 660  
QY 661 AACGTAAATTCGCGAGTACTGCTGATTTCCCAACGTACGGGCGCGCGGAGCTGATC 720  
DB 661 AACATGAATTTGCAAAATTAATCTGCATCATCCCTTATCTGGCCACGCGGCTGAGTGGT 720  
QY 721 GTGTTCTGCGCGCGCTGCTGCGCGCGGCTCGGCTTCTCTGTTCAACACCTATCCG 780  
DB 721 ATCTGCTGACGGCGATTGTCGGCGCGGCTGCGCTTCTCTGTTCAACACCTATCCG 780  
QY 781 CGCAGGTTCTCATGGCGAGCTCGCGCGGCTGCGCTGCGCGCGCGCTGCGGACCATC 840  
DB 781 CGCAGGTTCTTATGGGGGAGCTCGGTTCTCTGCTTGGCGCGCTGCGCTGCGGATATC 840  
QY 841 CGCGTGTGTCGCCAGGAGATGCTGCTGTTTATCATGTTGGTGGGCTGTTGCTCATGAA 900  
DB 841 GCGTACTGCTGCTGAGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900  
QY 901 ACCCTCTGGTGTGATCCAGTTCGCTTCTTCAAGTGAACCGACGCGCGCTCTTCGT 960  
DB 901 ACCCTGTCGTTGATTTTTCAGGTCGTTCTTCAAGCTGCGCGCTGCGGCTATCTCG 960  
QY 961 ATGGCGCGATCCATCCATTTTCGAAGCTGAAAGCTGGCGGACCGCGGCTGATGCTG 1020  
DB 961 ATGGCGCGATCCATCCATTTGAACTGAAAGCTGGCGGAGCGCGGCTGATGCTG 1020  
QY 1021 CGCTTCTGGATCATCACCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1080  
DB 1021 CGCTTCTGGATTTTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1080  
RESULT 7  
AAS53957  
ID AAS53957 standard; DNA; 1083 BP.  
XX AAS53957;  
AC AAS53957;  
DT 13-FEB-2002 (first entry)  
DE Klebsiella pneumoniae DNA for cellular proliferation protein #86.  
XX  
KW Antisense; ds; prokaryotic cellular proliferation gene; antibiotic;  
antibacterial; drug design.  
XX Klebsiella pneumoniae.  
OS WO200170955-A2.  
PN 27-SEP-2001.  
PD 21-MAR-2001; 2001WO-US009180.  
PF 21-MAR-2000; 2000US-0191078P.  
PR



PR 08-FEB-2002; 2002US-00072851.  
 PR 06-MAR-2002; 2002US-03626995.  
 XX PA  
 XX (ELIT-) ELITRA PHARM INC.  
 XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zvekind JW;  
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;  
 XX WPI; 2003-029926/02.  
 DR P-P8DB; ABU31662.  
 XX  
 XX New antisense nucleic acids, useful for identifying proteins or screening  
 PT for homologous nucleic acids required for cellular proliferation to  
 PT isolate candidate molecules for rational drug discovery programs.  
 XX  
 XX Claim 14; SEQ ID NO 23402; 1766pp; English.  
 XX  
 XX The invention relates to an isolated nucleic acid comprising any one of  
 CC the 6213 antisense sequences given in the specification where expression  
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:  
 CC (1) a vector comprising a promoter operably linked to the nucleic acid  
 CC encoding a polypeptide whose expression is inhibited by the antisense  
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated  
 CC polypeptide or its fragment whose expression is inhibited by the  
 CC antisense nucleic acid; (4) an antibody capable of specifically binding  
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular  
 CC proliferation or the activity of a gene in an operon required for  
 CC proliferation; (7) identifying a compound that influences the activity of  
 CC the gene product or that has an activity against a biological pathway  
 CC required for proliferation, or that inhibits cellular proliferation; (8)  
 CC identifying a gene required for cellular proliferation or the biological  
 CC pathway in which a proliferation-required gene or its gene product lies  
 CC or a gene on which the test compound that inhibits proliferation of an  
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a  
 CC compound's activity; (11) a culture comprising strains in which the gene  
 CC product is overexpressed or underexpressed; (12) determining the extent  
 CC to which each of the strains is present in a culture or collection of  
 CC strains; or (13) identifying the target of a compound that inhibits the  
 CC proliferation of an organism. The antisense nucleic acids are useful for  
 CC identifying proteins or screening for homologous nucleic acids required  
 CC for cellular proliferation to isolate candidate molecules for rational  
 CC drug discovery programs, or for screening homologous nucleic acids  
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,  
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target  
 CC prokaryotic essential genes. Note: The sequence data for this patent did  
 CC not form part of the printed specification, but was obtained in  
 CC electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 XX Sequence 1080 BP; 177 A; 302 C; 322 G; 279 T; 0 U; 0 Other;  
 Query Match 49.6%; Score 537.6; DB 7; Length 1080;  
 Best Local Similarity 68.6%; Pred. No. 3.5e-99;  
 Matches 741; Conservative 0; Mismatches 339; Indels 0; Gaps 0;  
 QY 1 ATGCTCTGCTGCTGGCGGATACCTGCAACAGTCTTACAGGGCTTCGGGCTTCCAG 60  
 DB 1 ATGTAGTATGCTGGCGGAACTTGGTCAATATATATTCGGCTTAAAGCTTTTCA 60  
 QY 61 TACGTACCTCGCGGCAATCTCAGCGTCTCACCGCGCTGCGCTGCTGCTGCTG 120  
 DB 61 TATCTGACGTTTCGGGCACTGTCAGCGCTGCTGACCGGCTGTCATCTGTTGGATG 120  
 QY 121 GGGCCCTGGATGATCGTACCTTGCAGATCCCCAGATCGGCCAGCGCGTGGCACGAC 180  
 DB 121 GGGCCCGGATGATCGCCCGCTGCAAAAACTCGCCTTTGGCCAGGTGCTAGTAAACGAC 180  
 QY 181 GGTCGCGAGTCGACCTGTGCAAGAGGGGACCCCGACCATGGCGGCGCCCTGATCCCTT 240  
 DB 181 GGCCCGGAGTCTGATTTTCAGTAAACGGGTACTCGACCATGGCGGATCATGATCCCTT 240  
 QY 241 ACCGCCATAGCCATCAGACGCTGCTGTGGGGGATCTTTCCAAACCGCTACGTGGGTA 300

Db 241 ACCGCCATAGCCATCAGACGCTGCTGTGGGGCTATCCATTAACCGTACGTCTGGTGC 300  
 QY 301 GTGCTGCTGCTTACCTGCTGTTTCGGTGCATTCGGCTGGGTAGACGACTACCCAGGTG 360  
 Db 301 GTACTACGGTATTAATTCGGCTACGGCATCATCGGTTTGGTGGATGATTACCGTAAAGTC 360  
 QY 361 ATCGAGAGAACTCCCGTGGCCTGCGAGCGCTGGAAGTACTTCTGGCAGTGGTGTTC 420  
 Db 361 GTGGCAAGATACCAAGGCTGATCGCCGCTGGAGTATTTCTGGATGTCGGTGATC 420  
 QY 421 GGCATCGCGCGCGCTGTTCTCTACATGATGCGGAACCCCGATCGAGACCACTG 480  
 Db 421 GCCTTGGCGCTGGCTTTCGGCTGATCTGCGGGGAAAGATACCCGCGCAACCGAGCTG 480  
 QY 481 ATCGTCCGATGCTGAAGAGCGCTCGAGATCCAGTGGGCACTTTCTTCGTGGTCTTGACC 540  
 Db 481 GTGGTGGCTTTTAAAGACGTAATGCCGACGTGGGCTGTTCTATATCTTCTGCTGCC 540  
 QY 541 TACTTCGTATCGCTGGCTCGAGCAATGCGGTGAACCTCAACCGAGTCTTCGAGCGCTG 600  
 Db 541 TACTTCGTATCGCTGGCTACCGCAACCGCTCAACCTGACCGAGCTTCGACGCGCTG 600  
 QY 601 GCGATCATCGGAGCGTAATGTTGCGGCGCGCTGGGCACTTCTGCTACCTGTCGGGC 660  
 Db 601 GCGATTCGCGACCGTTCGTCGCGGAGGCTTTCGCTGGTGGATGGGCCACCGGC 660  
 QY 661 AACCTGAAGTTCGCGAGTACCTGCTGATTCCCAACGTACCGGCGCGCGAGCTGATC 720  
 Db 661 AACATGAATTCGCCAACTACCTACATATCCGTTATCTGCGCAACGCGGAGCTGGT 720  
 QY 721 GTGTTTCGCGCGCTGGTGGCGCGGCTGGCTTCCTCTGTTCAACACCTATCCG 780  
 Db 721 ATCGTCTGTACGGCGATTTGTCGGGCGGGCTGGGCTTCTGTGTTCAACACCTATCCG 780  
 QY 781 GCGCAGGTCTTCATGCGGCGACGTCGCGCGCTGGCGCTGGGCGCGCGCTGGGCACCATC 840  
 Db 781 GCGCAAGTCTTATGCGGAGCTGCGTTCACTGCGCTGCGGCGCGCTGGGCATTATC 840  
 QY 841 GCGGTGATGTCGCGGCGGAGAGATCGTCTGTTTCAATGATGGTGGGTGTTGCTCATGAA 900  
 Db 841 GCGGTGCTGCTGCTCAGGAGTCTCTGCTGATGATGCGGCGGGTTCGTTGGTGGAA 900  
 QY 901 ACCCTCGGTGATGATCCAGTTCGCTTCAAGCTGACCGGACCGCGCTCTTCCT 960  
 Db 901 AGCTGCTGTTTATCTGAGTTCGCTTCTTTAAAGTTCGCGGTTCAGCGTATCTTCGC 960  
 QY 961 ATGGCGCGATPCCATCACCATTTCGAATGAAAGGCTGGCGGACCGCGCGGTGATGCTG 1020  
 Db 961 ATGGCGCGATPCCATCACCATTTCGAATGAAAGGCTGGCGGACCGCGCGGTGATGCTG 1020  
 QY 1021 CGCTTCGTGATCATCAGCTGATCTGCTGCTGATCGGCTTCGCACTTGAAGTGGCT 1080  
 Db 1021 CGCTTCGTGATCATCAGCTGATCTGCTGCTGATCGGCTTCGCACTTGAAGTGGCT 1080  
 XX  
 XX AC AC51132;  
 XX AC AC51132;  
 XX 19-JUN-2003 (first entry)  
 DE Prokaryotic essential gene #32789.  
 XX Antisense; ds; prokaryotic essential gene; cell proliferation;  
 KW drug design; gene.  
 XX Salmomella typhimurium.  
 XX WO20027183-A2.  
 XX 03-OCT-2002.  
 PD

RESULT 9  
 AC51132  
 ID AC51132 standard; DNA; 1083 BP.



DE Prokaryotic essential gene #33621.  
KW Antisense; ds; prokaryotic essential gene; cell proliferation;  
KW drug design; gene.  
XX  
XX Salmomella typhi.  
XX  
XX WO200277183-A2.  
XX  
XX 03-OCT-2002.  
XX  
XX 21-MAR-2002; 2002WO-US009107.  
XX  
XX 21-MAR-2001; 2001US-00815242.  
XX  
XX 06-SEP-2001; 2001US-00948993.  
XX  
XX 25-OCT-2001; 2001US-0342923P.  
XX  
XX 08-FEB-2002; 2002US-00072851.  
XX  
XX 06-MAR-2002; 2002US-0362699P.  
XX  
XX (ELIT-) ELITRA PHARM INC.  
XX  
XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;  
XX Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;  
XX P-PSDB; ABU48094.  
XX  
XX WPI; 2003-029926/02.  
XX  
XX P-PSDB; ABU48094.  
XX  
XX New antisense nucleic acids, useful for identifying proteins or screening  
XX for homologous nucleic acids, required for cellular proliferation to  
XX isolate candidate molecules for rational drug discovery programs.  
XX  
XX Claim 14; SEQ ID NO 39834; 1766pp; English.  
XX  
XX The invention relates to an isolated nucleic acid comprising any one of  
XX the 6213 antisense sequences given in the specification where expression  
XX of the nucleic acid inhibits proliferation of a cell. Also included are:  
XX (1) a vector comprising a promoter operably linked to the nucleic acid  
XX encoding a polypeptide whose expression is inhibited by the antisense  
XX nucleic acid; (2) a host cell containing the vector; (3) an isolated  
XX polypeptide or its fragment whose expression is inhibited by the  
XX antisense nucleic acid; (4) an antibody capable of specifically binding  
XX the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular  
XX proliferation or the activity of a gene in an operon required for  
XX proliferation; (7) identifying a compound that influences the activity of  
XX the gene product or that has an activity against a biological pathway  
XX required for proliferation, or that inhibits cellular proliferation; (8)  
XX identifying a gene required for cellular proliferation or the biological  
XX pathway in which a proliferation-required gene or its gene product lies  
XX or a gene on which the test compound that inhibits proliferation of an  
XX organism acts; (9) manufacturing an antibiotic; (10) profiling a  
XX compound's activity; (11) a culture comprising strains in which the gene  
XX product is overexpressed or underexpressed; (12) determining the extent  
XX to which each of the strains is present in a culture or collection of  
XX strains; or (13) identifying the target of a compound that inhibits the  
XX proliferation of an organism. The antisense nucleic acids are useful for  
XX identifying proteins or screening for homologous nucleic acids required  
XX for cellular proliferation to isolate candidate molecules for rational  
XX drug discovery programs, or for screening homologous nucleic acids  
XX required for proliferation in cells other than *S. aureus*, *S. typhimurium*,  
XX *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target  
XX prokaryotic essential genes. Note: The sequence data for this patent did  
XX not form part of the printed specification, but was obtained in  
XX electronic format directly from WIPO at  
XX ftp.wipo.int/pub/published\_pt\_sequences  
XX  
XX Sequence 1083 BP; 183 A; 262 C; 322 G; 316 T; 0 U; 0 Other;  
XX  
XX Query Match 47.8%; Score 518.2; DB 7; Length 1083;  
XX Best Local Similarity 67.4%; Pred. No. 2.9e-95;  
XX Matches 730; Conservative 0; Mismatches 353; Indels 0; Gaps 0;  
XX  
XX 1 ATGCTCTGCTGCGCGCAATACCTGCAACAGTTCTACAGGGCTTCGGCGTCTCCAG 60  
|||||  
|||||

1 ATGTTAGTTTGGCTGGCCGAGCATTTGGTCAAAATATTATTCGGGCTTTAACGTCCTTTCT 60  
61 TACCTGACCCCTGGCGGCAATTTCTCAGCGTGCTCACCGCGCTGCTGCTGCTGGGCTG 120  
61 TATCTGACGTTTGGCGCCATCGTCAGCCTGTTGACCGCGCTGTTCTATCTCTTTATGATG 120  
121 GGCCCTCGGATGATCCGCTACCTTTGAGATCCCGAGATCGGCCAGCCGCTGCGCAACGAC 180  
121 GGCCCGGATGATGCTGCTGCTGCAAAATCTCTTTGGCAGGTGCTAGCTAACGAT 180  
181 GGTCCGAGTCCGACCTGTCGAAGAGGCGACCCCGACCATGGCGGCGCCCTGATGCTT 240  
181 GGCCCGGATGCGCATTCAGTAAACCGGTAGCGCGACGATGGCGCGCATCATGATCTTG 240  
241 ACCGCCATAGCCATCAGACGCTGCTGTGGGGGATCTTTTCCAAACCGCTACGCTGGGTA 300  
241 ACCGCGATTTGATTTCCGTTCTGTTATGGGCTTACCGCTTAAACCGTACGCTCTGGTGC 300  
301 GTGCTGGTGTACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360  
301 GTGCTGGTGTATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360  
361 ATCGAGAGAACTCCCGTGGCTGCGGAGCGCGCTGGAAGTACTTCTGGCAGTCGGTGTTC 420  
361 GTGGGAAAGATACCAAGGGCTGATTGCGGCTGGAATATTTCTGGATGTCGGTTATC 420  
421 GGCATCGCGCCCGCTGTTCTCTCATGACTGCGCGAAACCCCGATCGAGACCACTG 480  
421 GCGCTCGCTGCGCTTGGGCTTTATCTGTCGGGAAAGACACGCGCGGACCCAACTG 480  
481 ATGCTGCGGATGCTGAAGAGCGTCCAGATCCAGTTCGGGCTCTTCTTCTGCTGCTGCTG 540  
481 GTGCTGCGCTTCTTTAAAGATGTTATGCCCAATTTGGGCTGTTTACATCTCTGCTGCTC 540  
541 TACTTCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600  
541 TACTTCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600  
601 GCGATCATGCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660  
601 GCGATCATGCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660  
661 AACGTGAAGTTTGGCGGAGTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720  
661 AACATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720  
721 GTGTTCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780  
721 ATTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780  
781 GCGCAGGTCTTTCATGGCGAGCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840  
781 GCGCAGGTCTTTCATGGCGAGCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840  
841 GCGGTGATGCTGCGCGAGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900  
841 GCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900  
901 ACCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960  
901 ACTCTG 960  
961 ATGGCGCGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020  
961 ATGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020  
1021 GCGTTCTGATCATCACCGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080  
1021 GCGTTCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1080  
1081 TGA 1083  
1081 TAA 1083

RESULT 11

AAS56340  
ID AAS56340 standard; DNA; 1083 BP.

12 XX AC .11350210 AAS56340;

DT 13-FEB-2002 (first entry)

DE *Salmonella typhi* DNA for cellular proliferation protein #373.

XX  
KW Antisense; ds; prokaryotic cellular proliferation gene; antibiotic;  
KW antibacterial; drug design.

OS salmonella typhi.

XX  
PN WO200170955-A2.

XX

PD 27-SEP-2001.  
XX  
PF 21-MAR-2001; 2001WO-US009180.

XX	21-MAR-2000;	2000US-01910
PR	23-MAY-2000;	2000US-02068
PR	26-MAY-2000;	2000US-02077
PR	23-OCT-2000;	2000US-02425
PR	27-NOV-2000;	2000US-02536
PR	22-DEC-2000;	2000US-02579
PR	16-FEB-2001;	2001US-02693
XX		
PA	(ELIT-) ELITRA PHARM INC.	

XX Haeselbeck R, Ohlisen KL, Zyskind JW, Wall D, Trawick JD, (

PI Yamamoto RT, Xu HH;

XX WPI; 2001-611495/70.

DR P-ESDB; AAU38481.

XX New polynucleotides for the identification and development of

PT antibiotics, comprise sequences of antisense nucleic acids.

PT

Claim 27; SEQ ID NO 9977; 511bp; English.

The invention relates to antisense inhibitors of genes essential to prokaryotic cellular proliferation, their use in identifying the genes, their use in the discovery of novel antibiotics, the essential genes themselves and the encoded proteins. The prokaryotes used are *Escherichia coli*, *Staphylococcus aureus*, *Salmonella typhi*, *Klebsiella pneumoniae*, *Pseudomonas aeruginosa* and *Enterococcus faecalis*. The invention is also useful for the identification of potential new targets for antibiotic development. The antisense nucleic acids can also be used to identify proteins used in proliferation, to express these proteins, and to obtain antibodies capable of binding to the expressed proteins. The proteins can be used to screen compounds in rational drug discovery programmes. The antisense nucleic acid sequence is also useful to screen for homologous nucleic acids which are required for cell proliferation in a wide variety of organisms. The present sequence encodes an essential prokaryotic cellular proliferation protein. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at [ftp.wipo.int/pub/published/pct\\_sequences](http://ftp.wipo.int/pub/published/pct_sequences)

Sequence 1083 BP; 183 A; 262 C; 321 G; 316 T; 0 U; 1 Other;

Query Match	47.8%	Score 517.2;	DB 4;	Length 1083;
Best Local Similarity	67.3%;	Pred. NO. 4.6e-95;		
Matches 723; Conservative	0;	Mismatches 354;	Indels	0;
Gaps	0;			

Qy 1 ATGCTCTGCTGTGGCCGAATACCTGCAACAGTTCTACAAGGGGCTTCGGCGTCTTCCAG 60

pB 1 ATGTTAGTTTGGTGGCCGAGCAATTTGGTCAAAATTATTCGGGGTTTACGTCTTTTCT 60

QY	61	TACCTGACCTCGCGGCAATCTCAGCGTGCTCACCGCGCTGCCTCTCGCTGCGGCTG	120
DB	61	TATCTGACCTTCGCGCCATCGTCAGCCTGTGTGACCGCGCTGTTCACTCTTTATAGATG	120
QY	121	GCGCCCTGATGATCCGTACCTTTGCAGATCCCCAGATCGGCCAGCGCTGCGCAACGAC	180
DB	121	GCGCCGCGTATGATCGCTCGTCTGCAGAAACTCTCTTTGGCCAGTGGTACGTAACGAT	180
QY	181	GGTCGCGAGTCGACCTGTCGAAGAAGGCGACCCGACCATGGGCGGCGCCTGATCCCT	240
DB	181	GCGCCGGAATCGACCTTCAGTAAACCGGCTACGCCGACGATGGGCGGCAATCATGATCCCTG	240
QY	241	ACCGCCATAGCCATCAGACCGCTGCTGTGGCGGATCTTTCCAAACCGCTACGTGTGGTGA	300
DB	241	ACGGCGATTGTATTTCCGTTCTGTATGGGCTTACCGCTCTAAACCGGTACGCTGTGGTGC	300
QY	301	GTGCTGGTTCGTTACCTGCTTCGGTGCCATCGCTGGTAGACGACTACCGCAAGGTG	360
DB	301	GTGCTGGTGATTTGATTCGGCTACGGCATATTCGTTTGTGCTGATGACTACGCCAAGTG	360
QY	361	ATCGAGAAGAACTCCCGTGGCTGCCGAGCGCTGGAAGTACTTTCTGGCAGTCCGTTGTC	420
DB	361	GTGCGGAAAGATACCAAAAGGCTGATTGGCGCTCGGAAATATTTCTGGATGTCGGTTATC	420
QY	421	GGCATCGGCGCGCGCTGTTTCTCTACATGACTGCCGAAACCCCGATCGAGACCAACCTG	480
DB	421	GCCTCGGCTGGCTTTGCGCTTATCTGTCGGGAAGACACGCCCGCAGACCCACTG	480
QY	481	ATCGTCCGATGCTGAAGAGCGTCGAGATCAGTGTGGGCAATCTTCTTGTGTCCTGACC	540
DB	481	GTGGTCCGCTCTTAAAGATGTTATGCCGCAATTTGGGCTGTTTTACATCTGCTGTCTC	540
QY	541	TACTTCGTATCGTCGGCTCGAGCAATCGGTGAACCTCACCGACGGTCTCGACGGCTG	600
DB	541	TACTTTGTATCGTCGTAGGGTAACGCGTAAACCTGACCGACGGGCTTGATGGTCTG	600
QY	601	CGCATCATGCCAGCGTAATGFTGCGCGCGCTGGGCAATCTTCTGCTACCTGTCCGGC	660
DB	601	GCATTAATGCCGACTGCTTCTGTTGCCGCGGCTTTGGCTGGTGGCTTTGGGCGACCGGG	660
QY	661	AAGTCGAAGTTCCCGAGTACCTGCTGATTCCCAACGTACCGGCGCGCGGAGCTGATC	720
DB	661	AAATGAACTTCGCCAATTACCTGCATATTCGGTATTTACGCCATCGCGGCGAGCTGGTG	720
QY	721	GTGTTTCGCGCGCTGTGTCGCGCGGCTCGGCTTCTCTGTTTCAACACCTATCCG	780
DB	721	ATTGCTGTACGCGCATTTGTCGCGCGGATTTAGGATTTCTTGTGTTTAAACCTATCCG	780
QY	781	GCCAGCTCTTCATGGCGAGCGTCGCGCGCTGGCGCTGGCGCGCGCTGGGCAACATC	840
DB	781	GCGCAGTTTATGGCGATGTCGATCGCTGGCGCTGGGCGCGCGTTGGCATTATC	840
QY	841	CGGCTGATCGTGCAGGAGATCGTGCTGTTTCATCATGGTGGGGTGTTCGTCATGAA	900
DB	841	CGGCTGCTGCTGGTTCAGGAAATTCGCTGGTATCATGGGCGGCTCTTTGTTGGTGGAA	900
QY	901	ACGCTCTCGGTGATGATCCAGGTCGCTTCTCTCAAGCTGACCGGACCGCGCTCTCCGCT	960
DB	901	ACTCTGCTCGTTCATCTCGAGTGGGTTCTTTAACTACGCGGACAGCGTATTTTCGCT	960
QY	961	ATGGCGCGGATCCATCAACATTTTCGAACCTGAAGAGCTGCGCGGACCGCGCTGATCGTG	1020
DB	961	ATGGCGCTATCCATCAACCACTAGAACTGAAGAGCTGCGCGGAAACCGCGCTGATGTTG	1020
QY	1021	CGCTTCGGAATCATACCGTGATCTCGTGTCTGATCGGCTCTCCGCCACTTTGAAGCTGCGT	1080
DB	1021	CGCTTCGGAATATTTTCGCTGATGCTGGTGCTGATTTGGCTTGGCAACGCTGAANGTACGT	1080
QY	1081	TGA	1083
DB	1081	TAA	1083

RESULT 13  
AAH81459

Qy 61 TACCTGACCCCTGCGCGGCAATCTCAGCGTGCTACCGCGCTGTCGCTGTCGCTGTCGCTG 120





RESULT 15  
ACA49179  
ID ACA49179 standard; DNA; 1060 BP.  
XX  
AC ACA49179;  
XX  
XX 19-JUN-2003 (first entry)  
XX  
DE DE  
XX  
XX Prokaryotic essential gene #30836.  
XX  
XX Antisense; ds; prokaryotic essential gene; cell proliferation;  
XX drug design; gene.  
XX  
XX Salmonella paratyphi.  
XX  
XX WO200277183-A2.  
XX  
XX 03-OCT-2002.  
XX  
XX 21-MAR-2002; 2002WO-US009107.  
XX  
XX 21-MAR-2001; 2001US-00815242.  
XX 06-SEP-2001; 2001US-00948993.  
XX 25-OCT-2001; 2001US-0342923P.  
XX 08-FEB-2002; 2002US-00072851.  
XX 06-MAR-2002; 2002US-0362699P.  
XX  
XX (ELIT-) ELITRA PHARM INC.  
XX  
XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen XL, Zyskind JW;  
XX Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;  
XX WPI; 2003-029926/02.  
XX P-PSDB; ABU45309.  
XX  
XX New antisense nucleic acids, useful for identifying proteins or screening  
XX for homologous nucleic acids required for cellular proliferation to  
XX PT isolate candidate molecules for rational drug discovery programs.  
XX  
XX Claim 14; SEQ ID NO 37049; 1766pp; English.  
XX  
XX The invention relates to an isolated nucleic acid comprising any one of  
XX the 613 antisense sequences given in the specification where expression  
XX of the nucleic acid inhibits proliferation of a cell. Also included are:  
XX (1) a vector comprising a promoter operably linked to the nucleic acid  
XX encoding a polypeptide whose expression is inhibited by the antisense  
XX nucleic acid; (2) a host cell containing the vector; (3) an isolated  
XX polypeptide or its fragment whose expression is inhibited by the  
XX antisense nucleic acid; (4) an antibody capable of specifically binding  
XX the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular  
XX proliferation or the activity of a gene in an operon required for  
XX the gene product or that has an activity against a biological pathway  
XX required for proliferation, or that inhibits cellular proliferation; (8)  
XX identifying a gene required for cellular proliferation or the biological  
XX pathway in which a proliferation-required gene or its gene product lies  
XX or a gene on which the test compound that inhibits proliferation of an  
XX organism acts; (9) manufacturing an antibiotic; (10) profiling a  
XX compound's activity; (11) a culture comprising strains in which the gene  
XX product is overexpressed or underexpressed; (12) determining the extent  
XX to which each of the strains is present in a culture or collection of  
XX strains; or (13) identifying the target of a compound that inhibits the  
XX proliferation of an organism. The antisense nucleic acids are useful for  
XX identifying proteins or screening for homologous nucleic acids required  
XX for cellular proliferation to isolate candidate molecules for rational  
XX drug discovery programs, or for screening homologous nucleic acids  
XX required for proliferation in cells other than *S. aureus*, *S. typhimurium*,  
XX *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target  
XX prokaryotic essential genes. Note: The sequence data for this patent did  
XX not form part of the printed specification, but was obtained in  
XX electronic format directly from WIPO at

CC ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 1060 BP; 180 A; 258 C; 313 G; 309 T; 0 U; 0 Other;  
Query Match 45.3%; Score 490.6; DB 7; Length 1060;  
Best Local Similarity 68.2%; Pred. No. 1.1e-89;  
Matches 710; Conservative 0; Mismatches 329; Indels 2; Gaps 2;  
QY 43 GGCCTCGCGTCTTCCAGTACCTGACCCCTGCGCGGCAATCTCAGCGTCTCACCGCGTG 102  
DB 19 GGCCTTAAACGCTCTTCTTATCTGACGTTTCGCGCCATCGTCAGCTGTTTACCGCGCTG 78  
QY 103 TCGCTGTCGCTGTGGCTGGGCGCTGATGATCGTACCTTGCAGATCCCGCCAGATCGGC 162  
DB 79 TTCACTCTTTATGATGGGCGCGGATGATGCTGCTGCAAAACTCTCTTTTGGC 138  
QY 163 CAGCGCTGCGCAACGACGCTCCGAGTCGACCTGTCGAAAGAGGCGACCCGACCATG 222  
DB 139 CAGGTGGTGACGTAACGATGGCGCGCAATCGACATTCAGTAAACGCGGTACGCCGACGATG 198  
QY 223 GGCGGCGCTGATCCCTTACCGCATAGCCATAGCAGCTGCTGCGGGGATCTTTC 282  
DB 199 GCGGCATCATGATCCTGACGCGGCAATGATTTCCGTTCTGTTATGGGCTTACCGTCT 258  
QY 283 AACCGTACGTGTGGGTAGTCTGCTGTACCTGCTGTTCCGTGCGCATCGGTGGGTA 342  
DB 259 AACCGTACGTGTGGGTAGTCTGCTGTGATTTGCTACGCGCATTTACGTTTGTGTC 318  
QY 343 GACGACTACCGCAAGGTGATCGAAGAGACTCCCGTGGCTGCCGAGCGCGTGAAGTAC 402  
DB 319 GATGACTACCAAGTGGTCCGAAAGATACCAAGGCGTGATTCGCGCGTGGAAATAT 378  
QY 403 TTCTGGCAGTCCGCTGTTCCGCGCATCGGCGCGCGCTGTTCTCTATCATGACTCCGAAACC 462  
DB 379 TTCTGGATGTCGGTATTCGCGCTCGG-GTGGCCCTTTCGCTTATTCGTCGGGAAGAC 437  
QY 463 CGATCGAGACACCCCTGATGCTGCCGATGCTGAAGAGCGTCGAGATCCAGTCCGATC 522  
DB 438 AGCGCGCGACCCCACTGGTGTGCCGCTCTTTAAAGATGTTATGCCGCAATTCGGGCTG 497  
QY 523 TTCTTGTGGTCCCTGACCTACTCTGTCATCGCTCGGCTCGAGCAATGGGTGAACCTCAC 582  
DB 498 TTTCATCTCTGCTGCTCTACTTTGTCATCGTCGCTAGCGGTACCGGTAAACCTGACC 557  
QY 583 GACGCTCGAGCGGCTGGCGATCATGCCGACGGTAATGTTGCGGCGCGCGCTGGGATC 642  
DB 558 GACGCGCTGTGATGCTGCGGATATGCGGACTGTTTTCGTTGCGCGCGGCTTTGCGCTG 617  
QY 643 TTCTGCTACCTGTGCGGCAAGTGAAGTTCCGCCAGTACCTGCTGATTCGCCAAGTACCG 702  
DB 618 GTGGCTGGGCGACCGGGAACATGAACTTCGCGCAATTCCTGATATTCGATTTAGCG 677  
QY 703 GCGCGCGCGAGCTGATCGTGTTCGCGCGCGCTGTTGCGCGCGCGGCTCGGCTTCCTC 762  
DB 678 TATGCGGCGAGCTGCTGATTTCTGTACGCGGATTTGTCGCGCGGGAATTAGGATTTTG 737  
QY 763 TGGTTCAACACTATCCGCGCGAGTCTTCATGCGCGAGCTGCGCGCGCTGCGCTGGGC 822  
DB 738 TGGTTTAAACATATCCGCGCGAGTCTTTTATGCGGAGCTGCGATCGCTGGCGTGGGC 797  
QY 823 GCGCGCTGGGACCATCGCGGTGATCGTGGCGCAGGAGATCGTCTGTTTCATCATGGGT 882  
DB 798 GCGCGTTGGGCAATTTATCGCGCTGCTGCTGCGTCAAGAGTTCTCTGCTGATCATGGC 857  
QY 883 GGGGTGTTGTCATGGAACCTCTCGGTGATGATCCAGGTGCGTCTCTTCAAGTACG 942  
DB 858 GCGGCTTTTGGTGGGAAACTCTGTGCGGTGATCTCTGCGAGTGGGTTCCTTTAAATACGC 917  
QY 943 GGACCGCGCGCTTTCGCTGATGGCGCGCATCCATCACCATTTCGAACTGAAAGGCTGGCG 1002  
DB 918 GGACAGC-TATTTTCGATGGCGCTTATCCATCACCACATGAACTGAAAGGCTGGCG 976  
QY 1003 GACCGCGCGGTGATCGTGGGCTTCTGGATCATACCGGTGATCTCTGCTGCTGATCGGCTC 1062

Db	977	GAACCGCGGTGATTGCGCTTCGGATTATTCGCTGATGCTGGTCTGATTGCCCTG	1036
Qy	1063	GCCACCTTGAAGCTGCGTTGA	1083
Db	1037	GCAACGCTGAAGGTACGTTAA	1057

Search completed: May 9, 2004, 13:38:24  
Job time : 495 secs

pred. No. is the number of results predicted by chance to have a

JOURNAL MEDLINE  
PUBMED  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
REMARK  
COMMENT  
FEATURES  
source

Pseudomonas aeruginosa: cloning, production, and purification of active enzymes  
Protein Expr. Purif. 21 (3), 393-400 (2001)  
21178826  
11281713  
2 (bases 1 to 5280)  
El-Sherbeini,M. and Azzolina,B.  
Pseudomonas aeruginosa mraY, murD, murE, and murF genes  
Unpublished  
3 (bases 1 to 5280)  
El-Sherbeini,M. and Azzolina,B.  
Direct Submission  
Submitted (20-SEP-2000) Biochemistry, Merck Research Laboratories,  
126 E. Lincoln Avenue, Rahway, NJ 07065, USA  
4 (bases 1 to 5280)  
El-Sherbeini,M. and Azzolina,B.  
Direct Submission  
Submitted (02-JAN-2001) Biochemistry, Merck Research Laboratories,  
126 E. Lincoln Avenue, Rahway, NJ 07065, USA  
Sequence update by submitter version replaced gi:10719753.  
On Jan 2, 2001 this sequence version replaced gi:10719753.  
Location/Qualifiers  
1..5280  
/organism="Pseudomonas aeruginosa"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:287"  
1..1464  
/gene="murE"  
1..1464  
/gene="murB"  
/note="UDP-N-acetylmuramyl-L-alanyl-D-glutamate:meso-2,6-diaminopimelate ligase"  
/codon\_start=1  
/transl\_table=11  
/product="MurE"  
/protein\_id="AAG45235.1"  
/db\_xref="GI:12007716"  
/translation="MPSLSQLPQAEEDLLIRELDSHGVRPVDLFLVPGHQDGR  
RAHIAADATKGTAVVAEAGELPPSDAPLIAVKGAAQSAVAGRFYGEPSRGID  
LIGVGTGKTSVSLVAQADLLGERGIVGTGTFYGALESGRITTPDLAVOAT  
LATLKOAGARAVAMEVSHGDDQGRVAGALGDIIVFTNLSRDHLDVHGSMEAYAAKA  
KLFAWPDRCRVINLDDDFRRRLAGEQDSLEIYSLTDSAFLYCREASPTASPTAR  
RSSIAKQDLSRPLGLRPNUSNLLAAVAGLLGYPGLDILRLTLPOLQPVGRMQLR  
GGGKPLVVDTPRETEASAAITADIRKGFADKVTFLPSRGEATLHIIASAAYDDVYL  
LAGKHEDYQIIGDVRHPPSDIEQAERAAAEVPHA"  
1457..2833  
/gene="murF"  
1457..2833  
/gene="murF"  
/note="recombinant enzyme catalyzed the ATP-dependent  
addition of D-alanine-D-alanine to the  
UDP-N-acetylmuramyl-L-alanine-D-glutamate-m-Dap precursor"  
/codon\_start=1  
/transl\_table=11  
/product="MurF"  
/protein\_id="AAG45236.1"  
/db\_xref="GI:12007717"  
/translation="MLEPLRLSQLTVALDARLIGEDAVFSAVSDTSRAIGPGLFIAL  
SGPRPDGHDVLAAYAAKGAVALVEREVADAPLQOLLVRDTAALGRIGALNRKFTG  
PLAAMTGGSGKTAVKEMLASILRTQAGDAESVLATRGNNLNDLGVPLTLQLAPOHRS  
AVIELGASRIGELAVTVELTRPHVALITNAGTAHVGEPPGPEKIVEAKGILLEGALD  
GTAVNLDDKDFATKARASRGLITFSLDRPQADFRADLQORARCGMFRLOQVAG  
EAQVQLNLDGRHNANALAAAHALGVPDGIIVAGLQALQPVKGRVAGLTAQSLR  
VIDGSYNANPASMALAIIDLSGFSGRITVLVLDMDGELGSAWAQAHREVGAYAAKQVSA  
CGSSEESH"  
2833..3915  
/gene="mraY"  
2833..3915  
/gene="mraY"  
/note="phospho-N-acetylmuramoyl-pentapeptide translocase"  
/codon\_start=1

/transl\_table=11  
/product="MurY"  
/protein\_id="AAG22121.1"  
/db\_xref="GI:10719754"  
/translation="MLLILAEYLOQFYKGFVQOYLTLRGLISLVLTLSLSIMLQPMW  
IRLQIPQIGAVRNDGQSHLSKGTQTYGKALILTAIAISTLMADLSNRYVYVVL  
VILLGAGWDDIRKXNRPGRKFWQSVFGIGAAVFLYMTATPIETTL  
IVPMKLSVEIQIGFIFFVITVIVGSSNANVLTDLGLAIMPVFMVAGALICYL  
SGNVFABYLLIPNVGAGELVFCALVAGLGLFWNTYPAQVFMGDVAGALGAA  
LGTIAIVRQEIIVLFIMGCVFMTLSVMIQVASFKLTCRRVFRMAPIHFFELKWP  
DPRVIVREWIITVILVLGLATLKLK"  
3929..5275  
/gene="murD"  
3929..5275  
/gene="murD"  
/note="UDP-N-acetylmuramyl-L-alanine:D-glutamate ligase;  
recombinant enzyme catalyzed the ATP-dependent  
addition of D-glutamate to the precursor sugar peptide"  
/codon\_start=1  
/transl\_table=11  
/product="MurD"  
/protein\_id="AAG45237.1"  
/db\_xref="GI:12007718"  
/translation="MSLIASDHFRIVVGLSGMSLVRYLARRGLPFAVVDTPRENPE  
LATRAQYQVEARCGELDAEFLCSARELYVSPGLSLRTPALVQAAGKVRISGIDIL  
FARBAKAPIVAITGSNAKSTVTLVGENAAVADKRVAVGNLGTALDLDADDILYV  
LELSSFOLETCDLNAEVATLVNVEDHMDVYDGMADYHLAKHRIERGARQVVRNADY  
ALTRLIADTVPCNSFGLNKDPFKAFGLIEEDGXWLAQFQKLLPVGELKIRGNHNY  
SNALAAALGHAVGLPFDMLGALAKFSLAHRKQWVRERQGVSTYDSSKATNVGAL  
AAIEGLGADIDGKLVLAGGDKGADFDHLEPFVARFCRAVVLRLDAGLIAQAALGNA  
VFLVRVATLDEAVRQAELAREGDAVLLSPACASLDNFKNFEERELFAKAVEBLA"  
Query Match 100.0%; Score 1083; DB 1; Length 5280;  
Best Local Similarity 100.0%; Pred. No. 1.5e-131; Indels 0; Gaps 0;  
Matches 1083; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 ATGTCCTCTGCTGCTGGCCGAATACCTGCAACAGTCTTCAAGGGCTTGGCGTCTTCAG 60  
Db 2833 ATGTCCTCTGCTGCTGGCCGAATACCTGCAACAGTCTTCAAGGGCTTGGCGTCTTCAG 2892  
Qy 61 TACCTGACCTGGCGGGCATCTTCAGCGCTCCACCGGCTTCGCTGCTGCTGGCTG 120  
Db 2893 TACCTGACCTGGCGGGCATCTTCAGCGCTCCACCGGCTTCGCTGCTGCTGGCTG 2952  
Qy 121 GGCGCCCTGGATGATCGTACCTTGCAGATCCCGCCAGATCGGCGAGCCGTGGCAACGAC 180  
Db 2953 GGCGCCCTGGATGATCGTACCTTGCAGATCCCGCCAGATCGGCGAGCCGTGGCAACGAC 3012  
Qy 181 GGTCCGACGTGGCACTGTGCAAGAGGCGACCCCGACCATGGCGGGCCCTGATCCTT 240  
Db 3013 GGTCCGACGTGGCACTGTGCAAGAGGCGACCCCGACCATGGCGGGCCCTGATCCTT 3072  
Qy 241 ACGCCATAGCATAGCACGCTGTGTGGCGGATCTTTCCAAACCGCTACGCTGGGTA 300  
Db 3073 ACGCCATAGCATAGCACGCTGTGTGGCGGATCTTTCCAAACCGCTACGCTGGGTA 3132  
Qy 301 GTGCTGCTGTTACCTGCTGTGTGGTCCATCGCTGGGTAGACGACTACCGCAAGGTG 360  
Db 3133 GTGCTGCTGTTACCTGCTGTGTGGTCCATCGCTGGGTAGACGACTACCGCAAGGTG 3192  
Qy 361 ATCGAAGAAGACTCCCGTGGCTGCCGAGCGCTGGAAGTACTTCTGGCAGTGGGTGTC 420  
Db 3193 ATCGAAGAAGACTCCCGTGGCTGCCGAGCGCTGGAAGTACTTCTGGCAGTGGGTGTC 3252  
Qy 421 GGCAATCGGCGCGCGCTGTTCCTTACATGACTGCCGAACCCCGATCGAGACCCCTG 480  
Db 3253 GGCAATCGGCGCGCGCTGTTCCTTACATGACTGCCGAACCCCGATCGAGACCCCTG 3312  
Qy 481 ATCGTCCGATCGTGAAGAGCGTCGAGATCCAGTGGGGATCTTCTGCTGCTGCTGACC 540  
Db 3313 ATCGTCCGATCGTGAAGAGCGTCGAGATCCAGTGGGGATCTTCTGCTGCTGCTGACC 3372

ORIGIN

QY	541	TACTTCGTCATCGTCGGCTCGAGCAATCGCGTGAACTCCTCAGCAGCGTCTCGAGCGGCTG	600
Db	3373	TACTTCGTCATCGTCGGCTCGAGCAATCGCGTGAACTCCTCAGCAGCGTCTCGAGCGGCTG	3432
QY	601	GGATCATGCCGACGATGATGCTGCGCGCGCTGGGCACTCTTCTGCTACTGTCGCGGC	660
Db	3433	GGATCATGCCGACGATGATGCTGCGCGCGCTGGGCACTCTTCTGCTACTGTCGCGGC	3492
QY	661	AACGTGAAGTTCCGCGAGTACCTGCTGATTCCTCCAACTACCGGCGCGCGGAGCTGATC	720
Db	3493	AACGTGAAGTTCCGCGAGTACCTGCTGATTCCTCCAACTACCGGCGCGCGGAGCTGATC	3552
QY	721	GTGTTCTGCGCGCGCTGCTGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	780
Db	3553	GTGTTCTGCGCGCGCTGCTGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	3612
QY	781	GGCGAGGTTTCATCGCGGACGTCGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCGC	840
Db	3613	GGCGAGGTTTCATCGCGGACGTCGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCGC	3672
QY	841	GGGTTGATCGTGGCGCAGGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	900
Db	3673	GGGTTGATCGTGGCGCAGGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	3732
QY	901	ACCTCTCGGTGATGATCCAGTCTGCTTCTTCAAGCTGACCGGACGCGGCTCTTCGGT	960
Db	3733	ACCTCTCGGTGATGATCCAGTCTGCTTCTTCAAGCTGACCGGACGCGGCTCTTCGGT	3792
QY	961	ATGGCGCGGATCATCACCATTTCGAAGTGAAGCTGCGCGGCGCGGCGGCGGCGGCGG	1020
Db	3793	ATGGCGCGGATCATCACCATTTCGAAGTGAAGCTGCGCGGCGCGGCGGCGGCGGCGG	3852
QY	1021	CGCTTCTGATCATCACCCTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1080
Db	3853	CGCTTCTGATCATCACCCTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	3912
QY	1081	TGA 1083	
Db	3913	TGA 3915	

  

RESULT 2	AE004856/c
LOCUS	AE004856
DEFINITION	Pseudomonas aeruginosa PAO1, section 417 of 529 of the complete genome.
ACCESSION	AE004856
VERSION	AE004856.1
KEYWORDS	GI:9950633
SOURCE	Pseudomonas aeruginosa PAO1
ORGANISM	Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales; Pseudomonadaceae; Pseudomonas.
REFERENCE	1 (bases 1 to 24000)
AUTHORS	Stover, C.K., Pham, X.-Q.T., Erwin, A.L., Mizoguchi, S.D., Warriner, P., Hickey, M.J., Brinkman, F.S.L., Hufnagle, W.O., Kowalik, D.J., Lagrou, M., Garber, R.B., Goltry, L., Tolentino, E., Westbrock-Wadman, S., Yuan, Y., Brody, L.L., Coulter, S.N., Folger, K.R., Kas, A., Larbig, K., Lim, R.M., Smith, K.A., Spencer, D.H., Wong, G.K.-S., Wu, Z. and Paulsen, I.T.
TITLE	Complete genome sequence of Pseudomonas aeruginosa PAO1, an opportunistic pathogen
JOURNAL	Nature 406 (6799), 959-964 (2000)
MEDLINE	20437337
PUBMED	10984043
REFERENCE	2 (bases 1 to 24000)
AUTHORS	Stover, C.K., Pham, X.-Q.T., Erwin, A.L., Mizoguchi, S.D., Warriner, P., Hickey, M.J., Brinkman, F.S.L., Hufnagle, W.O., Kowalik, D.J., Lagrou, M., Garber, R.B., Goltry, L., Tolentino, E., Westbrock-Wadman, S., Yuan, Y., Brody, L.L., Coulter, S.N., Folger, K.R., Kas, A., Larbig, K., Lim, R.M., Smith, K.A., Spencer, D.H., Wong, G.K.-S., Wu, Z., Paulsen, I.T., Reizer, J., Sailer, M.H., Hancock, R.E.W., Lory, S. and Olson, M.V.

TITLE	Direct Submission
JOURNAL	Submitted (16-MAY-2000) Department of Medicine and Genetics, University of Washington Genome Center, University Of Washington, Box 352145, Seattle, WA 98195, USA
REFERENCE	3 (bases 1 to 24000)
AUTHORS	Pseudomonas aeruginosa Community Annotation Project (PseudocAP)
CONSTRM	Direct Submission
JOURNAL	Submitted (04-FEB-2003) Department of Molecular Biology and Biochemistry, Simon Fraser University, 8888 University Dr., Burnaby, British Columbia V5A 1S6, Canada
COMMENT	----- This represents the February 3, 2003 version of the continually updated, reviewed, Pseudomonas aeruginosa PAO1 genome annotation, from PseudocAP (see <a href="http://www.pseudomonas.com">http://www.pseudomonas.com</a> for latest updates and links to alternate annotations). PseudocAP is coordinated by Fiona S.L. Brinkman (Simon Fraser University, Canada) and Robert E.W. Hancock (University of British Columbia, Canada). We welcome submission through <a href="http://www.pseudomonas.com">www.pseudomonas.com</a> of any proposed changes. ----- 'Protein name confidence' is used to rate our confidence of the accuracy of the protein name. Class 1: Function experimentally demonstrated in P. aeruginosa. Class 2: Function of highly similar gene experimentally demonstrated in another organism (and gene context consistent in terms of pathways its involved in, if known). Class 3: Function proposed based on presence of conserved amino acid motif, structural feature or limited sequence similarity to an experimentally studied gene. Class 4: Homologs of previously reported genes of unknown function, or no similarity to any previously reported sequences. -----
FEATURES	Location/Qualifiers
source	1. .24000 /organism="Pseudomonas aeruginosa PAO1" /mol_type="Genomic DNA" /strain="PAO1" /db_xref="taxon:208964" complement(91..1308) /genes="argJ" /locus_tag="PA4402" complement(91..1308) /genes="argJ" /locus_tag="PA4402" complement(91..1308) /genes="argJ" /locus_tag="PA4402" /notes="Protein name confidence: Class 2 (high similarity to functionally studied protein); Subcellular localization: Cytoplasmic; Subcellular localization confidence: Class 2; Alt protein name: ornithine transacetylase" /codon_start=1 /transl_table=1 /product="glutamate N-acetyltransferase" /protein_id="AAG07790.1" /db_xref="GI:9950634" /translation="MAVGIGPLSTLHPVPGFELGIASAGIKRPRKDVVVMRCAGGSSVAGVFLNAPCAAPVTLAKQRFELGEVRYLLTNGNNACTGEAGLAAGAATCTCAKAEALAGVATSVLPSTGIVGTEPLPVAKIKPMATLGYIADKVAQGLVQLLRDARPKGAHQFVHDGTVTVTIGISKGAMIKPMATLGYIADKVAQGLVQLLRDARPKGARITIDGTSNDCCMLIATGRAALPEVTSQALFAALQVLEVSMEIAQAVRDEGATKFTVTVNGGATGHECLDVGVAHSPSLIKTALFASDPNNGRIIAVGRAGVANL DYSKIDVFLGDCIASRGGRAASYTEQGAAMVAQAEIGIRIELGRGCTETITWTDLSHEYVKINAEYRT" complement(1455..4205) /genes="SecA" /locus_tag="PA4403" complement(1455..4205) /genes="SecA" /locus_tag="PA4403" /notes="Protein name confidence: Class 2 (high similarity to functionally studied protein)" /codon_start=1 /transl_table=1 /product="secretion protein SecA"
gene	
CDS	





/protein\_id="AAN66948.1"  
/db\_xref="GI:24982782"  
/translation="MMAAASRAGAAETOALEYLOGQLLARNWRCKGGELDIV  
MLDADTVFVVRVRLHAGFGALDSIDGRKQLVLAATLFLQKEAHWGNYPFRPDV  
VALQSHAGAPLQNLNAPEC"  
complement(2842..4659)  
/locus\_tag="PPI325"  
complement(2842..4659)  
/locus\_tag="PPI325"  
/codon\_start=1  
/transl\_table=11  
/product="lipoprotein, putative"  
/protein\_id="AAN66949.1"  
/db\_xref="GI:24982783"  
/translation="MIACLRLLTALCAALAAACASSFSSSLGELPRTPDASIEOLLE  
KAASKSADALLRLSAADLAYKQDFPRAARILEQVPLDTLKPAQOVFASTLAEL  
AMSNQPAALTAHLASLORLAELPADLOARTYSVHAALADGQALAAQORILIA  
PLLGKQHPDHPAALWALSLPAEQLOQPTNDOTLAGWTSLAFVAKSGATLQOQAA  
LEAMVQHPDHPAALWALSLPAEQLOQPTNDOTLAGWTSLAFVAKSGATLQOQAA  
AHFQAQSQGPAPAVQVDFSSRIGSLDFTFQAQAGVQLVIGLEPLRYKLAAPKQ  
LPITTLALNADAGQKAPQLFQGLAAEDAREVARRADGMVRAVALVPSPGEWD  
RVLAARFDWEGNGTLLAERIAQVALAQIADLPQLQSEGRAKSLQSTVGGSTA  
AQPRRQDIDFIFLASTPQAAQIKPTLNFQYAGDVPVYATSNLYSASGINVNDMN  
GIRFCTPWLDTNSLRQOVQWQPOAAGSLGLYANGVDVSLAPRLGLKALPDN  
RVGLGSLSLNANQVERQLPWAEEFSGGVKRLPDTVR"  
4837..5721  
/locus\_tag="PPI326"  
4837..5721  
/locus\_tag="PPI326"  
/note="identified by match to PFAM protein family HMM  
PF00590"  
/codon\_start=1  
/transl\_table=11  
/product="tetrahyrole methylase family protein"  
/protein\_id="AAN66950.1"  
/db\_xref="GI:24982784"  
/translation="MCAVDTPAGASKSTVGLYVATPIGNLDDMSARALKVADVAL  
IAAEDTRISRLQHFQDIDTFLAAACHEHNERDEGRFLSKLLAGENVALVSDAGTPII  
SDPQVHLVRQAQASVVPVPGACALIAALSAAGLPSDRFIFEGFLPAKQAGRRL  
FOVREPTLIFYAPAPRHLECLDEMEAFVGGDRPALRELTKTFTTLKGLPLAEUR  
AFVAGDSNQOGEQCVLVGWSAPEGEQAISAEQRVLDLILLALPLKRAAALAEIT  
GVKXNLLYQLALEXQKAE"  
complement(5744..6164)  
/locus\_tag="PPI327"  
/note="This region contains a gene with one or more  
premature stops or frameshifts, and is not the result of a  
sequencing artifact; conserved hypothetical protein,  
degenerate; similar to GB:X16944, and SP:P15488;  
identified by sequence similarity; putative"  
6455..6928  
/locus\_tag="PPI328"  
6455..6928  
/locus\_tag="PPI328"  
/note="identified by match to PFAM protein family HMM  
PF02381"  
/codon\_start=1  
/transl\_table=11  
/product="conserved hypothetical protein TIGR00242"  
/protein\_id="AAN66951.1"  
/db\_xref="GI:24982786"  
/translation="MWGITAVFRGANAVSLDAKGLRAMPYRDELDSRCNGOLIVTI  
DAVPCVCLVPIDWEQIIEAKLRALPSIRENRLQRLIGNAVDLIDGSGRFLVPP  
RLREYAKLDKXAMLVGQLKFLWEDDANNAVASADLAAIQQPGAMPDELRLIL"  
6925..7872  
/locus\_tag="PPI329"  
6925..7872  
/locus\_tag="PPI329"  
/note="similar to GP:9950652; identified by sequence  
similarity; putative"  
/codon\_start=1  
/transl\_table=11  
/product="conserved hypothetical protein TIGR00006"  
/protein\_id="AAN66952.1"

/db\_xref="GI:24982787"  
/translation="MTIDSGFNHTVLLDEAVEALALRADGCLYDITFGRGHSRLIL  
SKLGPQGLGDFDQPAIATGOALAEEDRFVIVORSFABLGAEVAARLHGKVSQV  
LIDLGVSSPOLDDPERGFSFLNDGPMNPDOQVSAEFTATAPVEIARVFKYEG  
ERFAGMARAVVEREQPTRTADLAELVKVANPEKGNPATRAFOGLRLHVN  
ELGDLAEGLAEALDALEVGRRLAVISFHSLEDRIYKLFMRKLIVKGEADNLPRNLPVQH  
KVFEPIKILIGKAQFASAEALKNPRSRKSVNRVAEKUR"  
7869..8162  
/gene="ftsL"  
/locus\_tag="PPI330"  
7869..8162  
/gene="ftsL"  
/locus\_tag="PPI330"  
/note="similar to GB:M11186, GB:M62611, SP:P01178,  
PID:189411, PID:386991, and PID:825699; identified by  
sequence similarity; putative"

	gene	CDS	Query Match	Best Local	Mismatches	Indels	Gaps
			75.7%; Score 820.2; DB 1; Length 301995;	Similarity	0;	0;	0;
			Matches 918; Conservative				
Qy	1	ATGCTCTCTCTGCTGGCCGAATACCTGCAACAGTTCTTCAAGGGCTTCGGCGTCTTCAG	60				
Db	12754	ATGCTCTCTCTGCTGGCCGAATACCTGCAACAGTTCTTCAAGGGCTTCGGCGTCTTCAG	12813				
Qy	61	TACCTGACCTGGCGGCAATCTCAGCGTGTCCAGCGCTGCTCGCTGCTGGCTG	120				
Db	12814	TACCTGACCTGGCGGCAATCTCAGCGTGTCCAGCGCTGCTCGCTGCTGGCTG	12873				
Qy	121	GGCCCTGGATGATCGTACCTTGCAGATCCCGCCAGATGGCCAGCGCTGCGCAACGAC	180				
Db	12874	GGCCCTGGATGATCGTACCTTGCAGATCCCGCCAGATGGCCAGCGCTGCGCAACGAC	12933				
Qy	181	GGTCCGACCTGCACCTGTGGAAGAGGCGACCCCGACCATGGCGGCGCCCTGATCCCT	240				
Db	12934	GGCCCGCAATCGACCTGTGCAAGTCGGGCAACCCGACCATGGGTGGGCGCTGATCCCT	12993				
Qy	241	ACCGCATAGCATCAGACGCTGCTGTGGCGGATCTTTTCAACCGCTACGTGTGGGTA	300				
Db	12994	TGGCCATCGCGTCTGAGCACTTGTGTGGCGGACCTGAGCAACCGCTATGTATGGGTT	13053				
Qy	301	GTGCTGGTGTACCTTGTCTTCCGTGCGATCGCTGGGTAGAGATACCGCAAGGTG	360				
Db	13054	GTGCTGATGTACCTTGTCTTCCGTGCGATCGCTGGGTAGAGATACCGCAAGGTG	13113				
Qy	361	ATCGAAGAACTTCCCGTGGCTTCCGAGCCGCTGGAAGTACTTCTGCGAGTCGGTGTTC	420				
Db	13114	ATCGAAGAACTTCCCGTGGCTTCCGAGCCGCTGGAAGTACTTCTGCGAGTCGGTGTTC	13173				
Qy	421	GGCATCGGCGCGCGCTGTTCCTCTACATGACTCGCGAAACCCCGATCGAGACCCCTG	480				
Db	13174	GGCATCGGCGCGCGCTGTTCCTCTACATGACTCGCGAAACCCCGATCGAGACCCCTG	13233				
Qy	481	ATCGTCCGATGCTGAAGAGCGTCCGAGATCCAGTTGGGATCTTCTTGGTCTGCTGACC	540				
Db	13234	ATCGTCCGATGCTGAAGAGCGTCCGAGATCCAGTTGGGATCTTCTTGGTCTGCTGACC	13293				
Qy	541	TACTTCGTATCGTGGCTCGAGCAATGGGTGAACCTCACCGAGGTCTCGACGGCCTG	600				
Db	13294	TACTTCGTATCGTGGCTCGAGCAATGGGTGAACCTCACCGAGGTCTCGACGGCCTG	13353				
Qy	601	GGCATATCGCGCGGTAAATGTTCCGCGCGCTGGGATCTTCTGCTACCTGTGCGGC	660				
Db	13354	GGCATATCGCGCGGTAAATGTTCCGCGCGCTGGGATCTTCTGCTACCTGTGCGGC	13413				
Qy	661	AACGTGAAGTTCGCGAGTACCTGCTGATTCACAGTACCGGGCGCGCGGAGTGTATC	720				
Db	13414	AACGTGAAGTTCGCGAGTACCTGCTGATTCACAGTACCGGGCGCGCGGAGTGTATC	13473				
Qy	721	GTGTTCTGCGCGCGCTGCTGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	780				
Db	13474	GTGTTCTGCGCGCGCTGCTGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	13533				

/protein\_id="AAN66948.1"  
/db\_xref="GI:24982782"  
/translation="MMAAASRAGAAETOALEYLOGQLLARNWRCKGGELDIV  
MLDADTVFVVRVRLHAGFGALDSIDGRKQLVLAATLFLQKEAHWGNYPFRPDV  
VALQSHAGAPLQNLNAPEC"  
complement(2842..4659)  
/locus\_tag="PPI325"  
complement(2842..4659)  
/locus\_tag="PPI325"  
/codon\_start=1  
/transl\_table=11  
/product="lipoprotein, putative"  
/protein\_id="AAN66949.1"  
/db\_xref="GI:24982783"  
/translation="MIACLRLLTALCAALAAACASSFSSSLGELPRTPDASIEOLLE  
KAASKSADALLRLSAADLAYKQDFPRAARILEQVPLDTLKPAQOVFASTLAEL  
AMSNQPAALTAHLASLORLAELPADLOARTYSVHAALADGQALAAQORILIA  
PLLGKQHPDHPAALWALSLPAEQLOQPTNDOTLAGWTSLAFVAKSGATLQOQAA  
LEAMVQHPDHPAALWALSLPAEQLOQPTNDOTLAGWTSLAFVAKSGATLQOQAA  
AHFQAQSQGPAPAVQVDFSSRIGSLDFTFQAQAGVQLVIGLEPLRYKLAAPKQ  
LPITTLALNADAGQKAPQLFQGLAAEDAREVARRADGMVRAVALVPSPGEWD  
RVLAARFDWEGNGTLLAERIAQVALAQIADLPQLQSEGRAKSLQSTVGGSTA  
AQPRRQDIDFIFLASTPQAAQIKPTLNFQYAGDVPVYATSNLYSASGINVNDMN  
GIRFCTPWLDTNSLRQOVQWQPOAAGSLGLYANGVDVSLAPRLGLKALPDN  
RVGLGSLSLNANQVERQLPWAEEFSGGVKRLPDTVR"  
4837..5721  
/locus\_tag="PPI326"  
4837..5721  
/locus\_tag="PPI326"  
/note="identified by match to PFAM protein family HMM  
PF00590"  
/codon\_start=1  
/transl\_table=11  
/product="tetrahyrole methylase family protein"  
/protein\_id="AAN66950.1"  
/db\_xref="GI:24982784"  
/translation="MCAVDTPAGASKSTVGLYVATPIGNLDDMSARALKVADVAL  
IAAEDTRISRLQHFQDIDTFLAAACHEHNERDEGRFLSKLLAGENVALVSDAGTPII  
SDPQVHLVRQAQASVVPVPGACALIAALSAAGLPSDRFIFEGFLPAKQAGRRL  
FOVREPTLIFYAPAPRHLECLDEMEAFVGGDRPALRELTKTFTTLKGLPLAEUR  
AFVAGDSNQOGEQCVLVGWSAPEGEQAISAEQRVLDLILLALPLKRAAALAEIT  
GVKXNLLYQLALEXQKAE"  
complement(5744..6164)  
/locus\_tag="PPI327"  
/note="This region contains a gene with one or more  
premature stops or frameshifts, and is not the result of a  
sequencing artifact; conserved hypothetical protein,  
degenerate; similar to GB:X16944, and SP:P15488;  
identified by sequence similarity; putative"  
6455..6928  
/locus\_tag="PPI328"  
6455..6928  
/locus\_tag="PPI328"  
/note="identified by match to PFAM protein family HMM  
PF02381"  
/codon\_start=1  
/transl\_table=11  
/product="conserved hypothetical protein TIGR00242"  
/protein\_id="AAN66951.1"  
/db\_xref="GI:24982786"  
/translation="MWGITAVFRGANAVSLDAKGLRAMPYRDELDSRCNGOLIVTI  
DAVPCVCLVPIDWEQIIEAKLRALPSIRENRLQRLIGNAVDLIDGSGRFLVPP  
RLREYAKLDKXAMLVGQLKFLWEDDANNAVASADLAAIQQPGAMPDELRLIL"  
6925..7872  
/locus\_tag="PPI329"  
6925..7872  
/locus\_tag="PPI329"  
/note="similar to GP:9950652; identified by sequence  
similarity; putative"  
/codon\_start=1  
/transl\_table=11  
/product="conserved hypothetical protein TIGR00006"  
/protein\_id="AAN66952.1"



/translation="MNAHFSLMWDYTBDELLGLIRRGVELKOLNRKGVLEPLKNRV  
LGMIFKSSNTEPLSPAGMQLGGQAIPLSHRDITOLGRCEPIADSAKVISRLMDAVM  
IRYAHNITFPAANSRVPVINCAGFAGIEFNEFEFLALAGERTVVRDPAKAAVAGHLVST  
NNCNSXIERAIGDFDQFVACVAGFAGIEFNEFEFLALAGERTVVRDPAKAAVAGHLVST  
DVTSMQCEETARKALFAPQVTRALLDLADKVDLVMHCLFAHRRGESEISVDLLDDA  
RSVAMDQENRLEHAQKALLEFLVAPCFQA"  
5026..6087  
/locus\_tag="PSPT04165"  
5026..6087  
/locus\_tag="PSPT04165"  
/note="similar to GB:Z48042, SP:Q14444, and PID:662994;  
identified by sequence similarity; putative"  
/codon\_start=1  
/transl\_table=11  
/product="ABC transporter, ATP-binding protein"  
/protein\_id="AAO57621.1"  
/db\_xref="GI:28854558"  
/translation="MLNLRNLACQYQNRVQDNLNLNAGDIGCLSSGGCKTIT  
LRAAGPEPIHAGETITLADGVISRPGLTAPKRRIGNVFQDIALFPHLSVADNIAPG  
IRQRPDLQVIGELLEVLNGLKRYPHLSGGQQORVALARALAPOLLILLDDPFF  
NLDEGLRRLSHVETIKARGTSAILVTHDQEEAFVSDHGVFKEGRLEQWDTFY  
SLNHEPOTPPFVSHVIGQYFIRGQMIPEPESVHTEGLVLRGNRAYPGVRGSAVDLLRP  
DDIVAPDSDLKARVIGTFKGAFTLYRLQLPTGSQLEAIFPSHADHRQEEVGIKVA  
ADHVLVFAAQGSVAAYAG"  
complement(6096..6566)  
/locus\_tag="PSPT04166"  
complement(6096..6566)  
/locus\_tag="PSPT04166"  
/note="similar to SP:P21484; identified by sequence  
similarity; putative"  
/codon\_start=1  
/transl\_table=11  
/product="Ybak/ebec protein"  
/protein\_id="AAO57622.1"  
/db\_xref="GI:28854559"  
/translation="MTPALDLKKNRABRHISYHDPKAAASYGLEAAEKLGLDAARV  
FKTLASTKEGELLVAVVPVVGITLDLKALANAAGAKTMDAPAAQRSTGYLLGGIS  
PLGQKRLRTFIDEPAQLFESIVSAGRRGLEVELTAVALAEHTQARFAPIGRE"  
5918..7775  
/gene="glpF"  
/locus\_tag="PSPT04167"  
5918..7775  
/gene="glpF"  
/locus\_tag="PSPT04167"  
/note="similar to SP:P11244; identified by sequence  
similarity; putative"  
/codon\_start=1  
/transl\_table=11  
/product="glycerol uptake facilitator protein"  
/protein\_id="AAO57623.1"  
/db\_xref="GI:28854560"  
/translation="WTIALKOPTLTGQVAFELGTALMIFFTGCVAAALKVAGATFGL  
WELICWNAVMGILYLSAGISGAHLNPAVSIALSLFAGFEKRLPFVISAQIAGFC  
GAGLVLYISLFFFEHNAHIIIRSESLASVFPNPAISVGAFLVEVYIT  
ILMGVIMALGDDNGLPRGALAPLLIGLLVAVIGSAMGPLTGFAFNPAIDFGPKLMTF  
FAGWGEIAFTGGRDIPYFLVPIFAPILGACLAGAAYRALIARHLPSVTIVNEKEAPV  
VRGKVQASS"  
7805..9310  
/gene="glpK"  
/locus\_tag="PSPT04168"  
7805..9310  
/gene="glpK"  
/locus\_tag="PSPT04168"  
/note="identified by match to PFAM protein family HMM  
PF02782"  
/codon\_start=1  
/transl\_table=11  
/product="glycerol kinase"  
/protein\_id="AAO57624.1"  
/db\_xref="GI:28854561"  
/translation="MTDQNKYIIALDQGTSSRAIFPRDANVVSTAQSEFVQHYP  
QAGVHDPMEIPATQACWTKALADLHDIQIAIGITNQRETIVTWERDTGRPIY  
NAIVWQCRSTEICKRDGLEEYIKDTTGLVLDLPYFSGSKVKWILDHVEGSRRAR

KGELMFGTITWLWIKFTGGKVVHVDYTNASRTMLFNIHTLEWQDRLMDLIDIPREIL  
PEVSSSEVYGHSSGIP IAGIAGDQQAALFGQMCVPGQAKNTYGTGCFLLMNTGKK  
AVKSAHMTLITGCGPRGEVAYALEGAVFNGSGTVQWLRLDELKILNLDLTFEYFAGKY  
KDSQXVLVPAFTGLGAVPWPYARGALFGLTRGVKVDHI TRAALES IAYOTRDLDA  
MQDSGERLKS LRYDGGAVANNFLMQFOADILGTHVSEPMQRETTALGAAFLAGLAIG  
FWSSUDEURNKAVIERVFEPAEEAHEKLYAGNQKAVARTEDWEPHENEE"  
9421..10176  
/gene="glpR"  
Query Match 70.9%; Score 767.4; DB 1; Length 311600;  
Best Local Similarity 81.9%; Pred. No. 6e-91;  
Matches 885; Conservative 0; Mismatches 196; Indels 0; Gaps 0;  
QY 1 ATGCTCTGCTGCTCGCGGCAATACCTGCAAGTGTCTACAGGGGCTTCGGGCTCTCCAG 60  
DB 292940 ATGCTGCTGCTGCTGCGCGAGTTCTTACAACAGTTCTTACAAGGCTTTTCGGGCTCTTCAG 292881  
QY 61 TACCTGACCTCGCGGGGATTCCTCAGCGTGTCTACCGGGCTGTCTCGCTGTCTGCTGTGGCTG 120  
DB 292880 TACCTGTCTCTCGCGGGGATTCCTTGGCGTGTCTCATCGCTGACGCTTTCGCTGTGCTG 292821  
QY 121 GGGCCCTGATGATCGGTACCTTGCAGATCCCGCCAGATCGCGGAGCGCTGGGCAACAC 180  
DB 292820 GGGCCCTGATGATCGGCACCTCGCAGATCGCGGAGCTCGGTTCGTTAACGAT 292761  
QY 181 GGTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240  
DB 292760 GGTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 292701  
QY 241 ACGCCATAGCCATCAGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300  
DB 292700 TGTCCATCGGTATCAGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 292641  
QY 301 GTGCTG 360  
DB 292640 GTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 292581  
QY 361 ATCGAAGAACTCCCGTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420  
DB 292580 ATTGAAAGAAATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 292521  
QY 421 GGCATCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480  
DB 292520 GGCCTTTCGCGCGCAATCTTCTGTCACGCTGCGCCATCGCGGACCGAACACCATTTG 292461  
QY 481 ATCGTCCGATGCTGAAGAGCGTCCAGATCCAGTTCGGCATCTTCTTCGTGCTGCTGCTGCTGCTG 540  
DB 292460 ATCGTCCGATGCTCAAGGATTCGCGATTCGCTGCGCATCGCTGCTGCTGCTGCTGCTGCTG 292401  
QY 541 TACTTCTG 600  
DB 292400 TACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 292341  
QY 601 GCGATCATCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660  
DB 292340 GCGATCATCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 292281  
QY 661 AACGTGAAGTTCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720  
DB 292280 AACGTGAAGTTCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 292221  
QY 721 GTGTTCTGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780  
DB 292220 GTCTTCTCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 292161  
QY 781 GCGCAGGTCTTCATGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840  
DB 292160 GCTCAGGTGTTTCATGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 292101  
QY 841 CGCGTGTGCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900  
DB 292100 GCGGTGTGCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 292041

QY	901	ACCTCTCGGTGATCCAGGTCGCTTCCTTCAAGCTGACCGGACGCGGCTCTTCGGT	960	source
Db	292040	ACCCCTTCACTGGTCATTCACTGGTGGCTTTCAAACTGACCGCGCGGTAATTCGC	291981	
QY	961	ATGGGCGCATCATCACCATTTCGATCGAAGAGCTGGCGGACCGCGGCTGATCGTG	1020	
Db	291980	ATGGGCGCATTCACCACTTCGATCGAAGAGCTGGCGGACCGGCTGATCGTG	291921	
QY	1021	CGCTTCTGATCATCACCCTGATCTCGTGGTCTGATCGGCTGCCACCTTGAAGCTGGT	1080	gene
Db	291920	CGTTTCTGATCATCACCCTGATCTCGTGGTCTGATCGGCTGCCACCTGGAAGCTGAGG	291861	CDS
QY	1081	T	1081	
Db	291860	T	291860	
RESULT 5				
AE012171				
LOCUS	AE012171	13045 bp	DNA	linear
DEFINITION	Xanthomonas campestris pv. campestris str. ATCC 33913, section 79			
ACCESSION	AE012171	AE008922		
VERSION	AE012171.1	GI:21111728		
KEYWORDS	Xanthomonas campestris pv. campestris str. ATCC 33913			
SOURCE	Xanthomonas campestris pv. campestris str. ATCC 33913			
ORGANISM	Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales; Xanthomonadaceae; Xanthomonas			
REFERENCE	1 (bases 1 to 13045)			
AUTHORS	da Silva, A.C.R., Ferro, J.A., Reinach, P.C., Farah, C.S., Furlan, L.R., Quaggio, R.B., Monteiro-Vitorello, C.B., Van Sluys, M.A., Almeida Jr., N.F., Alves, L.M.C., do Amaral, A.M., Bertolini, M.C., Camargo, L.E.A., Camarotte, G., Cannavan, F., Cardoso, J., Chamberg, F., Ciapina, L.P., Cicarelli, R.M.B., Coutinho, L.L., Cursino-Santos, J.R., El-Dorri, H., Faria, J.B., Ferreira, A.J.S., Ferreira, R.C.C., Ferro, M.I.T., Formighieri, E.F., Franco, M.C., Greggio, C.C., Gruber, A., Katsuyama, A.M., Kishi, L.T., Leite Jr., R.P., Lemos, E.G.M., Lemos, M.V.F., Locali, E.C., Machado, M.A., Madeira, A.M.B.N., Martinez-Rossi, N.M., Martins, E.C., Meidanis, J., Menck, C.F.M., Miyaki, C.Y., Moon, D.H., Moreira, L.M., Novo, M.T.M., Okura, V.K., Oliveira, M.C., Oliveira, V.R., Pereira Jr., H.A., Rossi, A., Sena, J.A.D., Silva, C., de Souza, R.F., Spinola, L.A.F., Takita, M.A., Tamura, R.E., Teixeira, E.C., Tezza, R.I.D., Trindade dos Santos, M., Truffi, D., Tsai, S.M., White, F.F., Setubal, J.C. and Kitajima, J.P.			
TITLE	Comparison of the genomes of two Xanthomonas pathogens with differing host specificities			
JOURNAL	Nature 417 (6887), 459-463 (2002)			
MEDLINE	22022145			
PUBMED	12024217			
REFERENCE	2 (bases 1 to 13045)			
AUTHORS	da Silva, A.C.R., Ferro, J.A., Reinach, P.C., Farah, C.S., Furlan, L.R., Quaggio, R.B., Monteiro-Vitorello, C.B., Van Sluys, M.A., Almeida Jr., N.F., Alves, L.M.C., do Amaral, A.M., Bertolini, M.C., Camargo, L.E.A., Camarotte, G., Cannavan, F., Cardoso, J., Chamberg, F., Ciapina, L.P., Cicarelli, R.M.B., Coutinho, L.L., Cursino-Santos, J.R., El-Dorri, H., Faria, J.B., Ferreira, A.J.S., Ferreira, R.C.C., Ferro, M.I.T., Formighieri, E.F., Franco, M.C., Greggio, C.C., Gruber, A., Katsuyama, A.M., Kishi, L.T., Leite Jr., R.P., Lemos, E.G.M., Lemos, M.V.F., Locali, E.C., Machado, M.A., Madeira, A.M.B.N., Martinez-Rossi, N.M., Martins, E.C., Meidanis, J., Menck, C.F.M., Miyaki, C.Y., Moon, D.H., Moreira, L.M., Novo, M.T.M., Okura, V.K., Oliveira, M.C., Oliveira, V.R., Pereira Jr., H.A., Rossi, A., Sena, J.A.D., Silva, C., de Souza, R.F., Spinola, L.A.F., Takita, M.A., Tamura, R.E., Teixeira, E.C., Tezza, R.I.D., Trindade dos Santos, M., Truffi, D., Tsai, S.M., White, F.F., Setubal, J.C. and Kitajima, J.P.			
TITLE	Direct Submission			
JOURNAL	Submitted (28-NOV-2001) Departamento de Bioquímica, Universidade de Sao Paulo, Av. Prof. Lineu Prestes 748, Sao Paulo, Sp 05508-900, Brazil			
FEATURES	Location/Qualifiers			

1. .13045	/organism="Xanthomonas campestris pv. campestris str. ATCC 33913"			
	/mol_type="genomic DNA"			
	/strain="ATCC 33913"			
	/db_xref="ATCC:33913"			
	/db_xref="taxon:190485"			
	/notes="pathovar: campestris"			
	189. .644			
	/gene="XCC0717"			
	189. .644			
	/gene="XCC0717"			
	/note="identified by sequence similarity; putative; ORF located using Blastx/Glimmer/Genemark"			
	/codon_start=1			
	/transl_table=11			
	/product="conserved hypothetical protein"			
	/protein_id="AA040032.1"			
	/db_xref="GI:21111729"			
	/translation="MHSVFOGETAITVDKGRMAVPTAYRDLVARVSGNRLVLTYPF EAGLWYAEKEWERYDDVMSKPTQVVRTLOQKLVSSAVLELDANGRLSIPASH RNAVGTEKKAVLLGMDKFEWSEQAHRALIQQTLSDGDLGDELLDLRL"			
	689. .1717			
	/gene="mraW"			
	/note="synonym: XCC0718"			
	689. .1717			
	/gene="mraW"			
	/note="identified by sequence similarity; putative; ORF located using Blastx/Glimmer/Genemark"			
	/codon_start=1			
	/transl_table=11			
	/product="cell division protein"			
	/protein_id="AA040033.1"			
	/db_xref="GI:21111730"			
	/translation="MSQPPAAHVPLVTVQVLEGLQVTEGTNGTLDGTGRGHARGVLE HLGGPDLVMDQPEQIAVAETHVETFGDARVSIRHGSFAGLGOVVAATVDGILLDLG VSSPQLDVAGRGFSFGDPLDMQDSCGSAAGLQAQTDREIADVLTGCEBROS RRTAATVAARGEQPLRLTQALADLASVMPRGDSKTHPATRSFQAIRHINRELAD EAGDAAIGALKPFGRLAVIFHSLEDRIYVKQFARYAKAPPNSRRULPEAQPFVPTLQ LVSGAIKADDSELAVNPRASAVLRVAELKGMENRESGKMGHGAASRPPTPSRFP TSPGNDAP"			
	1714. .1977			
	/gene="ftsL"			
	/note="synonym: XCC0719"			
	1714. .1977			
	/gene="ftsL"			
	/note="identified by sequence similarity; putative; ORF located using Blastx/Glimmer/Genemark"			
	/codon_start=1			
	/transl_table=11			
	/product="cell division protein"			
	/protein_id="AA040034.1"			
	/db_xref="GI:21111731"			
	/translation="MSRLILVLIACSIAGAIVYVYMHMRKLFVQLSKLEHSRDEL NIEGRQLQEQATWAESNRVDQVSRERIGMKFPETSDIVIRP"			
	2024. .3895			
	/gene="ftsI"			
	/note="synonym: XCC0720"			
	2024. .3895			
	/gene="ftsI"			
	/note="identified by sequence similarity; putative; ORF located using Blastx/Glimmer/Genemark"			
	/codon_start=1			
	/transl_table=11			
	/product="penicillin-binding protein 3"			
	/protein_id="AA040035.1"			
	/db_xref="GI:21111732"			
	/translation="WCAMKAGNRPRSNFNLGRLLVGVGLGCSVTLLIGRAVQL VNRDYOQGRARVLRPIATSRGMITDRNGEPLAVTPEVSIWNPEQLRPNDR I PELATGALPDLNKLAKQAGKEFMYLKRINPDKAHVLDLKI PQVFSORFRRP YFGQAMAHVLGFTNIDRGGQGL ELAFDEWLKRGSKGVTDARGAIVSSVDLVRP AQFGKDLTSLIDRIQFLAYELKRNALNVKAGSGSVTVMDVAGVETLWVNLPTYPD NSVTGINPSARNRAVTDLVDFPGTMTKPLTVATALTAGVVTGTDITIDNPGINAVGRF			

TIRDPVRNGLVNTGVTIRSSNTGAIAAKVDPQTFYQSIIRNFGYSGAPQGFPEE  
SSGVTLPQKWSGSKTMSYGLSVTPQIAQVATLNGSKLTPPTFVRQHNET  
RQVTPVEVAVIWMETVITQGGAKGAILGVHAGTGTARKNGAGGYERGHYNAL  
FAGLVPRPRFATVITNDPQGVYTGGLVSAPEVHRVMEGALRLMDVDPDIOSLW  
AAQAGKTQGVVPRQPADLDPALVDPVDEVSAGITFALASGATPAQAPLQESRQ"  
3892..5379  
/gene="murE"  
/note="synonym: XCC0721"  
3892..5379  
/gene="murE"  
/note="identified by sequence similarity; putative; ORF  
located using Blastx/Glimmer/Genemark"  
/codon\_start=1  
/transl\_table=11  
/product="UDP-N-acetylmuramoylalanyl-D-  
glutamate-2,6-diaminopimelate ligase"  
/protein\_id="AAM40036.1"  
/db\_xref="GI:21111733"  
/translation="MSRAMLSQLLPVALARDVQVSLVMSRAVRPQDAFVAIAGF  
GAHGFPAQALASGASAILPEPPAPADLPVADAIAPGLRGLRGLVLDHFGAPSQ  
NRMLGVGTNGKTSVQLLAQALTLTGTPGTIGTLGVGLYGNVFTGTTPLVLOT  
HAEALQADAGQAVMEVSQALDQGRVDAVQFVAVFNLTRDLDHGDMAQYGA  
AKALFAPAGARAVNLDLDAFGRTLFAALDPALHAIQVSSRAQAGATVQAQDLQD  
HGIQFTLHGAAPHSPLGRFNVNLLAVAGALHALDIAPAQIAEVLGRLOPQTHG  
RMRLLGAGAPLVVVDYAHPTDLEQALISLRGAQRILICVFCGGERDITGRPOM  
AAIAEVTADVAIVTDDNPRGEDGDIVIADILRGAEPDAAIVORDRAAAIQAISMA  
ADDIVLAGKHEPYQEVAGVYRHAFFDDAAVAQAALLPRTGLGRV"  
5376..6860  
/gene="murP"  
/note="synonym: XCC0722"  
5376..6860  
/gene="murP"  
/note="identified by sequence similarity; putative; ORF  
located using Blastx/Glimmer/Genemark"  
/codon\_start=1  
/transl\_table=11  
/product="UDP-N-acetylmuramoylalanyl-D-glutamyl-2,  
6-diaminopimelate-D-alanyl-D-alanyl ligase"  
/protein\_id="AAM40037.1"  
/db\_xref="GI:21111734"  
/translation="WKLPFLSLIAHFAGGELHGDAMI DAVGNDTRALASGLVVALR  
GERPDHDFADAVAGSGLLDRLIPALHVPQVAVNTALALARAQMDRATR  
VIALTSGNKTSTVALLAILAEAGRVGDTAYANPNRNEIGLPLAVIAEDDADL  
AVYENGAKPDGIAYLTDIAPDVALVNNIAPHLERMSLLGVAQPKGAIYIYALPAD  
GVAVINADDAQFEQLOAQPLQGRVLYRGLHASAEITARQLRHLAAQAQFVLAT  
PQGAPEVLYPGRHNLNLAATGLAGIALPIIAAGLAQATVAGRQIAHTLAT  
GAVLDDSYNPNQSLDAAIDALAAAPGEGWLVDNRELGPPEAPLHAAAGRRANA  
KLARIYALGALSAAQAQAFDQGRVFEATHAALAAALOADLAAASQTQDSSLMHEH  
TGAVARAASVAPPTLLVKGSRGSMDEIYVALLAPSEDASHAA"  
6850..7935  
/gene="mray"  
/note="synonym: XCC0723"  
6850..7935  
/gene="mray"  
/note="identified by sequence similarity; putative; ORF  
located using Blastx/Glimmer/Genemark"  
/codon\_start=1  
/transl\_table=11  
/product="phospho-N-acetylmuramoyl-pentapeptide-  
transferase"  
/protein\_id="AAM40038.1"  
/db\_xref="GI:21111735"  
/translation="WLLSLRWLQLESFLGNVLTFRGILALATLFLSLMWGPVAV  
TRKLAQFGKQPIRQDQPTQTHFSKAGTFTMGSSILITVLSLLWGLDNRNRYVWLVL  
AVMCFGAIWDDWIKVIRYDPPNGLKSRWYLLIQSIFGLAAGLFLYTDVPAITP  
YIPMFKALPLAGVSFVVIYFWIVGFSNANLTDGLDLGAINFTVLVACALGVAY  
ASGNVYFAYIKLPIPLAGBELIICSAIAGAGLGFPLWNTYPAWFMFGDIALSLGA  
VLGTVAIVTQRELVLVIMGVGVFVETLSVMTQVASFKLTKRVERMAPIHHLFKGM  
PEPRVIVRFWIIISVVLVLIGLATLKR"  
7935..9299  
/gene="ftsW"  
/note="synonym: XCC0724"  
7935..9299

Query Match 55.2%; Score 598; DB 1; Length 13045;  
Best Local Similarity 72.6%; Pred. No. 9.9e-69;  
Matches 788; Conservative 0; Mismatches 295; Indels 3; Gaps 1;  
1 ATGCTCTGCTGCTGGCGGAATACCTGCAAGAGTTCACAGAGGCTTCGGGGCTCTCCAG 60  
6850 ATGCTGCTTGAGCTGGCCCGCTTGGCTACAGCAACTGGAGAGCCTGTTGGGGCTGTCAAC 6909  
61 TACCTGACCTCGGGGCAATTCACAGCGTGTCTACCGGGCTGTCGCTGTGCTGTGGCTG 120  
6910 TATCTGACCTTCGGGGCAATCTGGCAGCGCTGACCGGGCTGTTCTCTGTGTTGGGATG 6969  
121 GGGCCCTGATGATCGGTACCTTGCAGATCCCGCAGATCGGGCAGCGCGTGGCAACGAC 180  
6970 GGGCCCGCGGTGATTTCGCAAGCTGGCGAGTTCAAGGGGGCGGCACCGGATTCGCCAGGAC 7029  
181 GGTCCCGACGTCCACCTGTGCAAGAAGGCAACCCCGACCATGGGGGGGCGCCCTGATCCTT 240  
7030 GGGCCGACAGCAATTTTTCGAAGCCGCGACCGACCATGGGTGGCTGCTGATCCTG 7089  
241 ACCGCCATAGCCATCAGACAGCTGCTGTGGGGCGGATCTTTCCAAACGGCTACGTGTGGTGA 300  
7090 CTACCGGTCACTTGTGCGTGTGCTGTGGGGCGGATCTGCGCAACCGCTATGTGTGGCTG 7149  
301 GTGCTGCTGTTACCTGCTGCTGTCGTCGTCATCGGCTGGGTAGACGACTACCGCAAGGTG 360  
7150 GTGCTGGCGGTGATGATCTGCTTCGGTGGATCGGTCGGTACGACGACTGGATCAAGATC 7209  
361 ATCGAAGAACTCCCGTGGCTGCGCGAGCCGCTGGAAAGTACTTCTGGCAGTCCGGTGTTC 420  
7210 GTGGCGCGCGCCCGAAACCGGCTTGAAGTCGCGCTGGAAATATCTGTCAGTCTGATCTTC 7269  
421 GGCATCGGCGCGCGCTGTTCTCTACATGACTGCGGAAACCCCGCATCGAGACCAACCTTCG 480  
7270 GGCCTGGCGCGCGGCTGTTCTCTGTAACACGCGGACGTCGCGCGCGGATCACCTTC 7329  
481 ATGCTGCCGATGCTGAAGACGCTCGAGATCCAGTTG---GGCATCTTCTTCGTGTCGTCG 537  
7330 TACATCCGATGTTCAAGGCCATCGCTTGCCTGCGCTGGCGGGGTGAGTTCGTGGTGTATC 7389  
538 ACTACTCTGTCATCGTCGGCTCGAGCAATGGGTGMACTCACCAGAGGCTCTCGACGCG 597  
7390 GCGTATTTCTGGATGTCGGCTCTTCCAAACGGGTCAACCTGACCGACGGCTTGACGCGC 7449  
598 CTGGCGATCATGCCGACGGTAAATGTTGTCGCGCGCGCTGGGCAATCTTCTGTACCTGTGCG 657  
7450 CTGGCGATCATGCCGACGGTGTGTCGCTGCGCTGCGCGCTGGGGGTCTTCTGCTATGCTCG 7509  
658 GGCACGTGAAGTTCGCGAGTACCTGCTGATTCCCAAGTACCGGCGCGCGCGCGGCGCTG 717  
7510 GGCACGTGGTATTTCGCGAATACCTCAAGATCCCGCTGATTCCCGGGCGCGGTGAGCTG 7569  
718 ATGCTGTTCTGGCGCGCGCTGCTGCGCGCGCGCTCGGCTTCTCTCTGTGTTCAACACCTAT 777  
7570 ATCATCATCTGCTCGCGGATTCGCGGGCAGGCTTGGGCTTTCGTGTGTTCAACACCTAC 7629  
778 CGGCGCAGGTCTTCAATGGGCGAGCTGCGCGCGCTGCGCGCTGGGGCGCGCGCTGGGCGAC 837  
7630 CGGCGCATGCTGTTCAATGGGCGACATCGCGCGGCTGTGCTGCGCGCGGTGCTCGGCGACC 7689  
838 ATCGCGGTGATCTGTCGCGCGCAGAGATCGTGTGTTTCATCATGGTGGGGGTGTTCTGTCATG 897  
7690 GTGGCGGTGATCTGTCGCTCAGGAACTGCTGCTGCTGATCATGCGCGCGGTGTTCTGTCATC 7749  
898 GAAACCTCTCGGTGATGATCCAGGTGCTGCTTCCATCAGCTGACCGGACCGCGCGCTCTTC 957  
7750 GAAACCTCTCGGTGATGATCCAGGTGCTGCGGCTGCGGCTGCAAGCTGACCGGCAAGCGCGTTC 7809  
958 CGTATGCGCGCGATCCATCACCATTTCGAATCGAAAGGCTGCGCGCGACCGCGCGCTGATC 1017  
7810 CGCATGCGCGCGCATCCATCCATCCCACTTCGAGCTCAAGGGCTGCGCGCGCGCGCGCTGATC 7869

```

QY      1018 GTGCGCTTCTGATCATACACCGTATCCCTGGTGTGATCGCGCTCGCCACCTTGAAGCTG 1077
Db      |||||||
7870 GTGCGCTTCTGATCATCTCGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 7929
QY      1078 CATTGA 1083
Db      |||||
7930 CCGCTGA 7935

RESULT 6
LOCUS   AE011708
DEFINITION Xanthomonas axonopodis pv. citri str. 306, linear BCT 29-MAY-2002
the complete genome.
ACCESSION AE011708
VERSION   AE011708
KEYWORDS  GI:21106891
SOURCE   Xanthomonas axonopodis pv. citri str. 306
ORGANISM Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
Xanthomonadaceae; Xanthomonas.
REFERENCE 1 (bases 1 to 13079)
AUTHORS da Silva,A.C.R., Ferro,J.A., Reinach,F.C., Farah,C.S., Furlan,L.R.,
Quaggio,R.B., Monteiro-Vitorello,C.B., Van Sluys,M.A., Almeida
Jr.,N.F., Alves,L.M.C., do Amaral,A.M., Bertolini,M.C.,
Camargo,L.E.A., Camarotte,G., Cannavan,F., Cardozo,J.,
Chambergo,F., Ciapina,L.P., Cicarelli,R.M.B., Coutinho,L.L.,
Cursino-Santos,J.R., El-Dorriy,H., Faria,J.B., Ferreira,A.J.S.,
Ferreira,R.C.C., Ferro,M.I.T., Formighieri,E.F., Franco,M.C.,
Greggio,C.C., Gruber,A., Katsuyama,A.M., Kishi,L.T., Leite
Jr.,R.P., Lemos,E.G.M., Lemos,M.V.F., Locali,E.C., Machado,M.A.,
Madeira,A.M.B.N., Martinez-Rossi,N.M., Martins,E.C., Meidanis,J.,
Mencck,C.F.M., Miyaki,C.Y., Moon,D.H., Moreira,L.M., Novo,M.T.M.,
Okura,V.K., Oliveira,M.C., Oliveira,V.R., Pereira Jr.,H.A.,
Rossi,A., Sena,J.A.D., Silva,C., de Souza,R.F., Spinola,L.A.F.,
Taktai,M.A., Tamura,R.E., Teixeira,E.C., Tezra,R.I.D., Trindade dos
Santos,M., Truffi,D., Tsai,S.M., White,F.F., Setubal,J.C. and
Kitajima,J.P.
TITLE Comparison of the genomes of two Xanthomonas pathogens with
differing host specificities
JOURNAL Nature 417 (6887), 459-463 (2002)
MEDLINE 22022145
PubMed 12042217
REFERENCE 2 (bases 1 to 13079)
AUTHORS da Silva,A.C.R., Ferro,J.A., Reinach,F.C., Farah,C.S., Furlan,L.R.,
Quaggio,R.B., Monteiro-Vitorello,C.B., Van Sluys,M.A., Almeida
Jr.,N.F., Alves,L.M.C., do Amaral,A.M., Bertolini,M.C.,
Camargo,L.E.A., Camarotte,G., Cannavan,F., Cardozo,J.,
Chambergo,F., Ciapina,L.P., Cicarelli,R.M.B., Coutinho,L.L.,
Cursino-Santos,J.R., El-Dorriy,H., Faria,J.B., Ferreira,A.J.S.,
Ferreira,R.C.C., Ferro,M.I.T., Formighieri,E.F., Franco,M.C.,
Greggio,C.C., Gruber,A., Katsuyama,A.M., Kishi,L.T., Leite
Jr.,R.P., Lemos,E.G.M., Lemos,M.V.F., Locali,E.C., Machado,M.A.,
Madeira,A.M.B.N., Martinez-Rossi,N.M., Martins,E.C., Meidanis,J.,
Mencck,C.F.M., Miyaki,C.Y., Moon,D.H., Moreira,L.M., Novo,M.T.M.,
Okura,V.K., Oliveira,M.C., Oliveira,V.R., Pereira Jr.,H.A.,
Rossi,A., Sena,J.A.D., Silva,C., de Souza,R.F., Spinola,L.A.F.,
Taktai,M.A., Tamura,R.E., Teixeira,E.C., Tezra,R.I.D., Trindade dos
Santos,M., Truffi,D., Tsai,S.M., White,F.F., Setubal,J.C. and
Kitajima,J.P.
TITLE Direct Submission
JOURNAL Submitted (28-NOV-2001) Departamento de Bioquimica, Universidade de
Sao Paulo, Av. Prof. Lineu Prestes 748, Sao Paulo, SP 05508-900,
Brazil
FEATURES
source
1. .13079
/organism="Xanthomonas axonopodis pv. citri str. 306"
/mol_type="genomic DNA"
/strain="306"
/db_xref="taxon:190486"
/notes="pathovar: citri"
193. .639
/gene="XAC0771"
CDS
193. .639
/gene="XAC0771"
/notes="identified by sequence similarity; putative; ORF
located using Blastx/Glimmer/Genemark"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AA035659.1"
/db_xref="GI:21106892"
/translations="MFQGETAITVDKGRMAVETAYRDLVAVSGNRLVLTYNPFAG
CLWLYAKEWVRDDVMKPTQVVRTLQKLVGSSAVLELDANGRLSIPASHRNA
VGTEKAVLLGMVDKFEWSEQAHRALIQQTLSDGLDGLLELLRL"
684. .1712
/gene="mrw"
/notes="synonym: XAC0772"
684. .1712
/gene="mrw"
/notes="identified by sequence similarity; putative; ORF
located using Blastx/Glimmer/Genemark"
/codon_start=1
/transl_table=11
/product="cell division protein"
/protein_id="AA035660.1"
/db_xref="GI:21106893"
/translations="MSQSPAARHVPVLYTVQLDGLQVTENGTYLDGTFGRGHARGVLE
HLGGGRLLVMDKDPRAIYVAHRSFGDARVSIHRGSAFAGLGOVAAATVDGILLDLG
VSPQDLVACRGSPFGKDGELMDMPDSQSAEFLAQSDEIADVLMTYGBEROS
RRARALVARRASQPLRLTRAQADLIASVMPGDSKTHPATRSFQIRIHINELADL
EAGLDALGALKFEGGLAVISFHSLSDRIVQFMARYAKAPPNNRRLPERAQPVALIQ
LVSGAKADDAELHVNPRARSAYLVAEKILGNGESGMKGNSAASRPFPTADSPFP
ASANGDAA"
1709. .1972
/gene="ftsI"
/notes="synonym: XAC0773"
1709. .1972
/gene="ftsI"
/notes="identified by sequence similarity; putative; ORF
located using Blastx/Glimmer/Genemark"
/codon_start=1
/transl_table=11
/product="cell division protein"
/protein_id="AA035661.1"
/db_xref="GI:21106894"
/translations="NSRLLLVLLACSIASIGVVMVNRHMRKLFVQLSKLEHTRDEL
NIFBGRQLQEQATWAESNRVDQVARARIGMKFETNDVVVRP"
2025. .3893
/gene="ftsI"
/notes="synonym: XAC0774"
2025. .3893
/gene="ftsI"
/notes="identified by sequence similarity; putative; ORF
located using Blastx/Glimmer/Genemark"
/codon_start=1
/transl_table=11
/product="penicillin-binding protein 3"
/protein_id="AA035662.1"
/db_xref="GI:21106895"
/translations="MKAGNRNPRSNFLRGLFLVGLALGCSVTIIGRAAYVOLVNR
DFYQRCGEARYLRELPATISRGMITDRNGEPLAVTSIVSNWPNQLLRNPDRIPEL
AKALGPLDLNLAQKAGKEFMYLQRINPDKAHAIVDLKI PGVFSQREFFRFPQ
GEAMHVLGFTNIDDRGOEGLAFDEWLRKPGSKVVDARGAIVESVDLVPAQP
KGLTISIDRIQFLAYKELRNALVENNAAGSMVMDVATGEVLAMVNLVYNPSV
TGTLNPSARNRNVATDLVEFGSTMKPLTAVALTAGVVTXDTVIDTTPGYMSVGRFTIR
DVPNRNGLNVTVITRSNIGAKTAAPVDQTFYQSIINFGYGAHSFGFCESAG
VILQPARWSGSKTMSYGLSVLPQIAQAYATGLNGKLTPTTFYVRGQHNTRQV
ITTEVARQVIMMETVYTGGAAGAAI LGHVAGKTGTARKNGANGYERGHYNALFAG
LTPATPRFATVIVNDPQKVGVLGVPFHRVMEGALRLMDVPDDIQSWLAAQ
AAGKTGPVVKPQSDPDPALVPDVEVSAGIPTALAPAPAAAQAPAPLQSRQ"
3890. .5377
/gene="murZ"
/notes="synonym: XAC0775"
3890. .5377
/gene="murZ"

```



```

RESULT 7
AE016925/c
LOCUS
DEFINITION
Chromobacterium violaceum ATCC 12472 section 16 of 16 of the
complete genome.
ACCESSION
AE016925 AE016825
VERSION
AE016925.1 GI:34105514
KEYWORDS
SOURCE
ORGANISM
Chromobacterium violaceum ATCC 12472
Chromobacterium violaceum ATCC 12472
Bacterii; Proteobacteria; Betaproteobacteria; Neisseriales;
Neisseriaceae; Chromobacterium.
1 (bases 1 to 208524)
.
Brazilian National Genome Project Consortium
The complete genome sequence of Chromobacterium violaceum reveals
remarkable and exploitable bacterial adaptability
Proc. Natl. Acad. Sci. U.S.A. 100 (20), 11660-11665 (2003)
22892880
14500782
2 (bases 1 to 208524)
Vasconcelos, A.T.R., de Almeida, D.F., Almeida, F.C., de
Almeida, L.G.P., de Almeida, R., Gonçalves, J.A.A., Andrade, E.M.,
Antonio, R.V., Araujo, J., de Araujo, M.F.F., Filho, S.A., Azevedo, V.,
Battista, A.J., Batista, L.A.M., Batista, J.S., Belo, A., vander Berg, C.,
Blaney, J., Bogo, M., Bonato, S., Bordignon, J., Brito, C.A.,
Brocchi, M., Burity, H.A., Camargo, A.A., Cardoso, D.D.P.,
Carneiro, N.P., Carraro, D.M., Carvalho, C.M.B., Cascado, J.C.M.,
Cavada, B.S., Chueire, L.M.C., Pasa, T.B.C., Duran, N., Fagundes, N.,
Falcao, C.L., Fantinatti, F., Farias, I.P., Felipe, M.S.S.,
Ferrari, L.P., Ferro, J.A., Ferro, M.I.T., Franco, G.R., de
Fretas, N.S.A., Furlan, L.R., Gazzinelli, R.T., Gomes, E.A.,
Gonçalves, P.R., Grangeiro, T.B., Grattapaglia, D., Grisard, E.C.,
Guimaraes, C.T., Hanna, E.S., Hungria, M., Jardim, S.N., Laurino, J.,
Leoi, L.C.T., Fassarella, L., Lima, A., Loureiro, M.F., Lyra, M.C.P.,
Macedo, M., Madeira, H.M.F., Mantio, G.P., Maranhão, A.Q.,
Martins, W.S., di Mauro, S.M.Z., de Medeiros, S.R.B., Meissner, R.V.,
Menck, C.F.M., Moreira, M.A.M., Nascimento, P.F., Nicolas, M.F.,
Oliveira, J.G., Oliveira, S.C., Paixão, R.F.C., Parente, J.A.,
Pedrosa, F.O., Pena, S.J.D., Ferreira, J.O., Pereira, M.,
Pinto, L.S.R.C., Pinto, L.S., Porto, J.I.R., Potrich, D.P.,
Neto, C.E.B., Reis, A.M.M., Rigó, L.U., Rondinelli, E., dos
Santos, E.B.P., Santos, F.R., Schneider, M.P.C., Seunarez, H.N.,
Silva, A.M.R., da Silva, A.L.C., Silva, D.W., Silva, R., Simoes, I.C.,
Simon, D., Soares, C.M.A., Soares, R.B.A., Souza, E.M., Souza, K.R.L.,
Souza, R.C., Steffens, M.B.R., Steindel, M., Teixeira, S.R.,
Urmenyi, T., Vettore, A., Wassem, R., Zaha, A. and Simpson, A.J.G.
Direct Submission
Submitted (22-JAN-2003) Labinfo, LMCC - Laboratório Nacional de
Computação Científica, Rua Getúlio Vargas 333, Petropolis, RJ
25651070, Brazil
FEATURES
Location/Qualifiers
1..208524
/organism="Chromobacterium violaceum ATCC 12472"
/mol_type="genomic DNA"
/strain="ATCC 12472"
/db_xref="taxon:243365"
complement(96..563)
/gene="pili2"
/locus_tag="CV4209"
complement(96..563)
/gene="pili2"
/locus_tag="CV4209"
/notes="identified by sequence similarity; putative; ORF
located using GeneMark/Blastx/COG2165"
/codon_start=1
/transl_table=11
/product="type IV pilin"
/protein_id="AAQ61869.1"
/db_xref="GI:34105515"
/translation="MKQKQKQFTLIELMIVVAIVGLTAATAIPAYODYTKRQVSE
GLAADAATAVEYATNNAWASGTAPNTTYGLAASITGNSVSNVLANGVQIA
YNSTVSNGLALVDLVPTVSAGSVQWYCTYTGSTGTTAIAASQLSANWVSTCRQ"
838..2184
/gene="bioA"
/locus_tag="CV4210"
838..2184
/gene="bioA"
/locus_tag="CV4210"
/EC_number="2.6.1.62"
/notes="identified by sequence similarity; putative; ORF
located using GeneMark/Blastx/COG0161"
/codon_start=1
/transl_table=11
/product="adenosylmethionine-8-amino-7-oxononanoate
transaminase"
/protein_id="AAQ61870.1"
/db_xref="GI:34105516"
/translation="MSNOALERSFAAVHPCTQMKERLPIVPIASAGVWLTD
FKYLDGVSSWNWNLFGCHPRIKQAVRKQDSLEHVMLAGFTHRPVLSERLAALS
GLHAFYSGDASATEIALKQSFHYKNGVQPKTRFVLSNYSHTGTGALAVDVP
LFSTYAEILKPGLRAPSPDAPVGRADVALRAARLENLLAKHGGEIAVIVE
PLVQGAAGMAMHDPVILTELRLCDQYQVHLIADEIAVGFRTSGSFAYQQAGITP
LCLSKGITGGFLTLCVLTREIYQAFYHDDVARGFLHSHSYTGNALACAAALVDI
FDDEDVLEANKVKAEGFTROMRPLARPEIRHFRQGMIAWFDVDTARSDFALAYESA
MLNEGCLLRPIGKSVFMPFPTLKSEMDWLASVSDALDEVLAGESVGGDPTALP"
2347..3816
/locus_tag="CV4211"
/locus_tag="CV4211"
/locus_tag="CV4211"
/notes="identified by sequence similarity; putative; ORF
located using Glimmer/Blastx"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAQ61871.1"
/db_xref="GI:34105517"
/translation="MROARPGVLPAGPRTDGSRAALQAQALSEGAPLLSAA
GERS
PARRRGLAONAAAGHQRQAVQVPGLOANPARKILPFAAGATQAQAGAAVHDD
GGALPHQAPAKLLSARLLAGLPANVRAAGARLAGVAGREHAGRLHLPASAR
SDHQAPVAPRPAIGQATGLRQATGATVAGAGSDGPGCLGAGPGCGPAALSAG
IAPVRLPARTAGRDRTQGTGTAQDGGRRLLSGAERSPAEPFSGMGMAFPT
QACPGIGDGGDHRRRHLPQRQLAARRRRPRRQRPAPARLLPADCRQPE
RIWLPAPASRAALRRGAVAGPAPAGVSVRSALGIDALRRQSGMRRRDAGP
AAAAGEGHAQHPQRQPTGLAAARQEPAGAAVAPSLQPVARVQASGRKRATG
PRFOASAIPFPFADGPPQRAILTWKRRAFSVTLPSL"
3845..4963
/gene="proB"
/locus_tag="CV4212"
3845..4963
/gene="proB"
/locus_tag="CV4212"
/EC_number="2.7.2.11"
/notes="identified by sequence similarity; putative; ORF
located using Glimmer/Genemark/Blastx/COG0263"
/codon_start=1
/transl_table=11
/product="glutamate 5-kinase"
/protein_id="AAQ61872.1"
/db_xref="GI:34105518"
/translation="MHSVTHASDRIVVVCSSLVNDGCLDNLALAAEVEALKE
RGQVVLVSSCAIAEGCQRLGWTVPKGVHQLQAAVAGQGLCAQYESAPRFGILT
AQILLTHEDADRTYLNARSTVTLNLNVVFIINENDVATSEIRFGNDITGALV
TNLEADALVLTDRGLYSADPRKPDFAEFHQAEAGDERLEDMAGGASVGTGGM
ITKILAAKRAASGAATVYASGREPHVLSLADGEAIGTQLVATNRMAARKQLADH
KLKAGRLDDGGAALAIRESGTSLLPVGSVAVGDFLRGACVADAAAGHEVARGLVN
YSSDEARQIVRKSTIREAALGYIVPELHHRDNVAL"
5057..6130
/gene="potF1"
/locus_tag="CV4213"
5057..6130
/gene="potF1"
/locus_tag="CV4213"
/notes="identified by sequence similarity; putative; ORF
located using

```



KEYWORDS

SOURCE Salmonella typhimurium LT2  
ORGANISM Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales; Enterobacteriaceae; Salmonella.  
REFERENCE 1 (bases 1 to 22348)  
AUTHORS McClelland, M., Sanderson, K.E., Spieth, J., Clifton, S.W., Latreille, P., Courtney, L., Porwollik, S., Ali, J., Dante, M., Du, F., Hsu, S., Layman, D., Leonard, S., Nguyen, C., Scott, K., Holmes, A., Grew, S., Mulvaney, E., Ryan, E., Sun, H., Florea, L., Miller, W., Stoneking, T., Nhan, M., Waterston, R. and Wilson, R.K.  
TITLE Complete genome sequence of Salmonella enterica serovar Typhimurium LT2  
JOURNAL Nature 413 (6858), 852-856 (2001)  
MEDLINE 21534948  
PUBMED 11677609  
REFERENCE 2 (bases 1 to 22348)  
AUTHORS The Salmonella typhimurium Genome Sequencing Project  
CONSTRM Direct Submission  
JOURNAL Submitted (29-MAR-2001) Genome Sequencing Center, Department of Genetics, Washington University School of Medicine, 4444 Forest Park Boulevard, St. Louis, MO 63108, USA  
COMMENT Supported by NIH grant 5U 01 AI43283

Coding sequences below are predicted from manually evaluated computer analysis, using similarity information and the programs; GLIMMER; <http://www.tigr.org/softlab/glimmer/glimmer.html> and GeneMark; <http://opal.biology.gatech.edu/GeneMark/>  
RC numbers were kindly provided by Junko Yabuzaki and the Kyoto Encyclopedia of Genes and Genomes; <http://www.genome.ad.jp/kegg/>, and Pedro Romero and Peter Karp at EcoCyc; <http://ecocyc.PangeaSystems.com/ecocyc/>

The analyses of ribosome binding sites and promoter binding sites were kindly provided by Heladia Salgado, Julio Collado-Vides and ReguonDB; [http://kinich.cifn.unam.mx:8850/db/regulondb\\_intro.frameset](http://kinich.cifn.unam.mx:8850/db/regulondb_intro.frameset)

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistries or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one m13 subclone.

FEATURES

source Location/Qualifiers  
1..22348  
/organism="Salmonella typhimurium LT2"  
/mol\_type="genomic DNA"  
/strain="LT2; SGSC 1412; ATCC 700720"  
/db\_xref="ATCC:700720"  
/db\_xref="taxon:99287"  
/notes="LT2"  
complement(298..915)  
/genes="leuD"  
/notes="synonym: STM0110"  
complement(298..903)  
/genes="leuD"  
/EC\_number="4.2.1.33"  
/notes="3-isopropylmalate dehydratase small subunit.  
(SW:LEUD\_SALTY)"  
/codon\_start=1  
/transl\_table=11  
/product="3-isopropylmalate isomerase (dehydratase), subunit with Leu"  
/protein\_id="AAL19074.1"  
/db\_xref="GI:16418609"  
/translation="MAKTLVYKFLDPAHVPVPEAPNETPLLYIDRLHVEHTVSPQAFDGL RAHHRVPVROGKTFATMDHNVSTQTKDINASGEMARLOMQLIKNCNEFGVELYDLNH PYQIVHVMGPEQGVLPFGMTIVCGDSHTAIFGAFALAPGIGTSVEHVVALSQTILKQ GRATKMKIEVTGNAAGITAKDILVLAIIIGKTSAGGTGHVVEFCGDAIRALSMEGRMT LCNMAEMGAKAGLVAPDETTFYVYGRHLHAPKGRDFDEAVEYWKLTITDDGATFDTV VALRABEIAQVTVGTNPGQIVSVTQIIIPDPASFPDVERASAEKALAYMGLQPGVPL TDVAIDKVFISCTNRIEDIRAAAEVAKGRKVAQVQALVVPFGSPVKAQAAGELD KIFIEAGFEWRLPGCSMCLAMNDRLNPGERCASITNRNPEGRQGRGRTHLYSPAMA AAAATGCFADIRSIK"  
complement(2317..3422)  
/genes="leuB"  
/notes="synonym: STM0112"  
complement(2317..3408)  
/genes="leuB"  
/EC\_number="1.1.1.85"  
/notes="3-isopropylmalate dehydrogenase. (SW:LEU3\_SALTY)"  
/codon\_start=1  
/transl\_table=11  
/product="3-isopropylmalate dehydrogenase"  
/protein\_id="AAL19076.1"  
/db\_xref="GI:16418611"  
/translation="MSKNYHIAVLPGDGGIGPEVMAQALKVMDAVRSEFDMRIITSHVD VGIADNHHGHPFKATVECEQADAILFGSGVGPKNENLPPEQPERGALLPLRKHFKLFSNURPAKLYQLEAFCEADIAANGFDILCVRELTGGIYFPGQREGSGGQYEHK AFTEVYHFEIERIAFIAPESARKRKVTSIDKANVLOSSLIMREIVNDVAKTYPD VELAHYVINDATMLQIKDSQFDVLLCSNLFGLIDECAMITSGMGLPFSASINQOG FLYVEPAGSGAPDIAGNTANPIAQILSLALLRLYSILDANDAATAIEQAINRALREGV RTGDLARGAAVSTDEMGDIARYVAEGV"  
complement(2322..2327)  
/genes="leuB"  
/notes="putative RBS for leuC; RegulonDB:STW51H000491"  
complement(3408..4999)  
/genes="leuA"  
/notes="synonym: STM0113"  
complement(3408..4979)  
/genes="leuA"  
/EC\_number="4.1.3.12"  
/notes="similar to E. coli 2-isopropylmalate synthase (AAC73185.1); Blastp hit to AAC73185.1 (523 aa), 92% identity in aa 1 - 523"  
/codon\_start=1  
/transl\_table=11  
/product="2-isopropylmalate synthase"  
/protein\_id="AAL19077.1"  
/db\_xref="GI:16418612"  
/translation="MSQQVIFDITLRDGEALQASLAKELQIALALERMGMVDVME VGFVSSPGDFESVQTARTIKNSVCALARCVEKDIDVAQAQKLVADAFRIHTFIAT SPWHIATKLSTLDEIVRAVYVVKRARNYTDPEFSCBDAQRTPVDDLARVVAEAIN AGARTNIPDTVGYTWPFAGIISGLYERVNDIKAIISVTHDDGLAVGNSLAHV HAGAQVEGAMNGIGERAGNCALEEVIMALKVKDKIMVYHNHINHWRTSQTVSQI CNWPIIPANKAIIUSGAFASSSGHQDGVLNKREYELMTPEISGLNQILNLSRSGR AAVKREMGYKDTYNDMDHLYADLKLADKKGQVDFDLKALAFINKQBEPEHFR LDYFSQGGSSDIATASVKLACGEIKAEAGATNDGPVDAIQAINITGYDVELVKYD LNFQKQKGLQGVDIWNHGRFRFHVGVLTDIVESAKAMVHVLNINWRAEVEKE LQKQAKNKNKETY"  
complement(3417..3422)  
/genes="leuB"

RBS

```
/note="putative RBS for leuB; RegulonDB:STMS1H0000492"
complement(4994..4999)
/gene="leuA"
/note="putative RBS for leuA; RegulonDB:STMS1H0000493"
complement(5067..5298)
/gene="leuU"
/note="synonym: STM0114"
complement(5067..5153)
/gene="leuU"
/note="leu operon leader peptide. (SW:LPL_SALTY)"
/codon_start=1
/transl_table=11
/product="leu operon leader peptide"
/protein_id="AAL19078.1"
/db_xref="GI:16418613"
/translation="MSHIVRETGLLLNAFIVGRPVGGIQH"
complement(5162..5167)
/gene="leuU"
/note="putative RBS for leuU; RegulonDB:STMS1H0000494"
complement(5211..5279)
/gene="leuU"
/note="putative -10 signal for leuU;
RegulonDB:STMLTH004472"
complement(5290..5298)
/gene="leuU"
/note="putative -35 signal for leuU;
RegulonDB:STMLTH004472"
5798..5754
/gene="leuU"
/note="synonym: STM0115"
5798..5803
/gene="leuU"
/note="putative RBS for leuU; RegulonDB:STMS1H0000495"
5810..5754
/gene="leuU"
/note="probable activator protein in leuabc operon.
(SW:LEUO_SALTY)"
/codon_start=1
/transl_table=11
/product="putative LysR family transcriptional regulator"
/protein_id="AAL19079.1"
/db_xref="GI:16418614"
/translation="MPEVTEKPHLLDMGPKQLRMVDNLTLTFDAVMQSNITRAAH
TLGMSQPAVNAVLKMFNDELFRVYGRGIQIOTARAFQIFVSRQALQVQNELPG
SGFPTSERVNLVCSPLNDILTSIYNRVEKIAPIFVFKASINQTEHQLRYO
ETEFSVIEEERPEFTSPVLFKDEMLVARSKEPRI:SGPLLEGDVNEOHAVUSDJR
YASPSQPNYDTPDKQSSVAYOGMALISLVNVVQSOTHLVATPRWLAEPEAESLDLQIL
PLPLKLSRTCYLSWHEAAGRDQKHQWMDLLVSVCKR"
6941..6953
protein_bind
Query Match 47.8%; Score 518.2; DB 1; Length 22348;
Best Local Similarity 67.4%; Pred. No. 2e-58;
Matches 730; Conservative 0; Mismatches 353; Indels 0; Gaps 0;
QY 1 ATGCTCCTGCTGCGCGAATACCTGCAACAGTCTACAGGGCTTCGGGCTTCCAG 60
Db 17556 ATGTTAGTTGGCTGGCGAGATTGTGTCAAATATTATTCGGCTTTAACGCTTTTCT 17615
QY 61 TACCTGACCTGCGCGGATTCTCAGCGTGTCCAGCGCTGTGCTGCTGCTGCGTG 120
Db 17616 TATCTGACGTTTCGGCCATGCTCAGCTGTGACCGGCTGTTCACTCTTTATGATG 17675
QY 121 GGGCCTGATGATCGTACCTTCAGATCCCCAGATCGCGCAGCGCGCTGCGCAAGCAC 180
Db 17676 GCGCCGCGATGATCGCTGCTGTCGAAAACCTCTTTTGGCCAGGTGGTACGTAACGAT 17735
QY 181 GGTCCGCGAGTCGACCTGTCGAGAAAGGCGACCCGACCATGGCGCGGCGCTCATCTT 240
Db 17736 GCGCCGCGATCGCATCTTCAGTAAACGCGGTACGCGAGATGGCGGATCATGATCTG 17795
QY 241 ACCGCCATGACCATCAGCAGCTGCTGTGGCGGATCTTTCAACCGCTACGTTGGGTA 300
Db 17796 ACGTCGATTGTAITTCGCTTCTGTATGGGTTACCGCTCTAACCCGATACGTCGTGTC 17855
```

```
QY 301 GTGCTGGTGGTACCTCTGTTCCGTTGCCATCGCTGGTAGACGACTACCGAAGTG 360
Db 17856 GTGCTGGTGGTATTTGATCGGCTACCGCAATATTCGGTTTTGTGATGACTACCGCAAGTG 17915
QY 361 ATCGAAGAAGAACTCCCGTGGCCCTGCGAGCCGCTGGAAGTACTTCTGGCAGTCGGTGTTC 420
Db 17916 GTGCGGAAAGATACCAAGGGCTGATTGCGCGCTGGAATATTTCTGGATGTCGGTTATC 17975
QY 421 GGCATCGGCGCGCGCTGTTCTCTACATGATCGCGAAACCCCGATCGAGACCACTG 480
Db 17976 GCGCTCGGCTGGCCCTTCGCGCTTTATCTCTCGGGAAGACACCGCGGACCAACTG 18035
QY 481 ATCGTCCGATGCTGAAGAGCGTCGAGATCCAGTTGGGCATCTTCTTGGTCTCTGACC 540
Db 18036 GTGGTGCCGCTCTTTAAAGATGTTATCGCGCAATGGGCGCTGTTTACATCTGCTGTC 18095
QY 541 TACTTCTGATCGTGGCTCGAGCAATGGGTGAACCTCACGAGCGGTCTCGACGCCCTG 600
Db 18096 TACTTTGTCTATCGTGGTACGGGTAAACCGGTGAACCTGACCGACGGGCTTGTGCTG 18155
QY 601 GGCATCATCGGACGGTAAATGTTGCGCGCGCGCTGCGCATCTTCTGTACTCTGCGGC 660
Db 18156 GCGATTATCGGACTGCTCTCGTTGCGCGCGGCTTTCGCTGGTGGCTTGGCGGACCGG 18215
QY 661 AAGTGAAGTTCCGCGAGTACCTGCTGATTCACCACTGACCGGCGCGCGGAGCTGATC 720
Db 18216 AACATGAACCTTCGCAATTAACCTGCATATTCGATATTTACGCCATGCGCGGCGAGCTGGT 18275
QY 721 GTGTTCTGCGCGCGCTGGTTCGCGCGCGCTTCGCTTCCTCTGCTTCAACACCTTATCCG 780
Db 18276 ATTGCTGTACGCGGATGTTGCGCGCGGATAGGATTTCTTGTGTTTAAACCTTATCCG 18335
QY 781 GGCAGGCTTTCATGGCGGACGTCGCGCGCTGCGCTGCGGCGCGCGCTGGGACCATC 840
Db 18336 GGCAGGTTTTTATGGCGGATGTCGCGATCGCTGGCGCTGGGCGGCTGCGCTGATTC 18395
QY 841 GCGGTGATCGTCGCGCAGGAGATCGTGTGTTTCATGCGGTGGGTGCTTCGTCTATGAA 900
Db 18396 GCGGTGCTGCTGCTGCGGATTTCTGCTGTGATCATGGCGGCGCTTTTGTGCTGAA 18455
QY 901 ACCCTCTCGTATGATCCAGTCGCTTCCTTCAAGCTGACCGGACCGCGCTCTTCGCT 960
Db 18456 ACTCTGCTGCTATCTCTGCGGTGGTTCCTTTAAACTACGCGGACGAGCTATTTTCGT 18515
QY 961 ATGGCGCGATCCATCACCATTTCGAAGCTGCGGACCGCGCGCTGATCGTG 1020
Db 18516 ATGGCGCGCTATCCATCACCCTATGAACTGAAAGCTGCGCGGAAACCGCGCGGTG 18575
QY 1021 CGCTTCTGGATCATCACCGTGTATCCTGTGCTGATCGGCTCGCCACCTTGAAGCTGCT 1080
Db 18576 CGCTTCTGGATTTATTCGCTGATGCTGTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTG 18635
QY 1081 TGA 1083
Db 18636 TAA 18638
```

```
RESULT 9
LOCUS AL627265
DEFINITION Salmonella enterica serovar Typhi (Salmonella typhi) strain CT16,
complete chromosome, segment 1/20.
ACCESSION AL627265 AL513382
VERSION AL627265.1 GI:16501283
KEYWORDS
SOURCE Salmonella enterica subsp. enterica serovar Typhi
ORGANISM Salmonella enterica subsp. enterica serovar Typhi
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
Enterobacteriaceae; Salmonella.
REFERENCE 1 (bases 1 to 251050)
AUTHORS Parkhill, J., Dougan, G., James, K.D., Thomson, N.R., Pickard, D.,
Wain, J., Churcher, C., Mungall, K.L., Bentley, S.D., Holden, M.T.G.,
```

Sebahia,M., Baker,S., Basham,D., Brooks,K., Chillingworth,T., Connerton,P., Cronin,A., Davis,P., Davies,R.M., Dowd,L., White,N., Farrar,J., Feltwell,T., Hamlin,N., Haque,A., Hien,T.T., Holroyd,S., Jagels,K., Krogh,A., Larsen,T.S., Leather,S., Moule,S., O'Gaora,P., Parry,C., Quail,M., Rutherford,K., Simmonds,M., Skelton,J., Stevens,K., Whitehead,S. and Barrell,B.G.

Complete genome sequence of a multiple drug resistant Salmonella enterica serovar Typhi Cn18

Nature 413 (6858), 848-852 (2001)

21534947

11677608

2 (bases 1 to 251050)

Parkhill,J.

Direct Submission

Submitted (25-OCT-2001) Submitted on behalf of the Salmonella sequencing team, Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA, UK

E-mail: parkhill@sanger.ac.uk

Notes:

Details of S. typhi sequencing at the Sanger Centre are available on the World Wide Web.

(URL, [http://www.sanger.ac.uk/Projects/S\\_typhi/](http://www.sanger.ac.uk/Projects/S_typhi/)).

Location/Qualifiers

1..251050

/organism="Salmonella enterica subsp. enterica serovar Typhi"

/mol\_type="genomic DNA"

/strain="Ct18"

/db\_xref="taxon:90370"

190..255

/gene="STY0001"

/note="synonym: thrL"

190..255

/gene="STY0001"

/note="Orthologue of E. coli thrL (LPT\_ECOLI); Pasta hit to LPT\_ECOLI (21 aa), 86% identity in 21 aa overlap"

/codon\_start=1

/transl\_table=11

/product="thr operon leader peptide"

/protein\_id="CAD01154.1"

/db\_xref="GI:16501284"

/db\_xref="SPTREMBL:Q8XG12"

/translation="MNRISTTITITITITGNAG"

337..2799

/gene="STY0002"

/note="synonym: thrA"

337..2799

/gene="STY0002"

/note="Orthologue of E. coli thrA (AKIH\_ECOLI); Pasta hit to AKIH\_ECOLI (820 aa), 94% identity in 820 aa overlap"

/codon\_start=1

/transl\_table=11

/product="aspartokinase I/homoserine dehydrogenase I"

/protein\_id="CAD01155.1"

/db\_xref="GI:16501285"

/db\_xref="GOA:Q8Z9R7"

/db\_xref="SPTREMBL:Q8Z9R7"

/translation="NRVTKFGTGVANAEFLRVADILLESNRQGVATVUSAPAKIT NRVAMITGGQDALPNI SDAERIFSDLLAGLASAQPGFLRLKRVVVEQFAQIK HVLGIILGQGDPSINAAI LCRGKMSIATMAGLLEARGHRVTIDPEVKELVAGHY LESTVDIAESTRRRIAASQIPADHMLMAGFTAGNEKGLVVLGRNGSDYSAAVLAACL RADCCIEIWTVDGVGTYCDPROVPDARILKMSYQEMELS YFGAKVLHPRTITPIAOF QPCLIKNTGNPQAGPTLIGASSDDNLPVKGISLNNMAMFVSVSGPMKGMIGMAR VFAPMSKISVLLITQSSSEYSISFCVPSQDCARARRAQDEBYELKEGLLEPLAV TERLAIISVGDGMRILGISAFAALARANINIVAQAQSSERSISVVNNDDAT GRVTHQMFNDQVIEFVIGVGVGGALLIEQLKQOTWLNKHI DLRCVGNVANSKA LLTNVHGLNLDNQABLAQANAPNLGRLIRLVKEHYLLNFVIVDCTSSQAVADQYAD FLREGHVVTNKNKSMIDYTHQLRPAQAQRKFLYDTNVGAGLPVIENTQLNLN AGDELQKFSGLISGLSIFIGLECMSLSQATALAREMGYTEPDPRLDLSGMDVARK LLILARETGRELSDIVIEPVLDPEDFASGDVTAFAHLPLQDDAFAARVAKARDG KVLRVGNIEEDGVCVKIAEVDGNDPLFPKVGKNGENALAFISHTYQPLPLVIRGYGAG NDVTAAGVPADLLRLTSLKWLIG"

337..1101

misc\_feature

/gene="STY0002"

/note="Pfam match to entry PF00696 aakinas, Amino acid kinase family, score 207.70, E-value 1.7e-58"

343..369

/gene="STY0002"

/note="PS00324 Aspartokinase signature"

1819..2781

/gene="STY0002"

/note="Pfam match to entry PF00742 Homoserine dh, Homoserine dehydrogenase, score 535.60, E-value 3.4e-157"

2314..2382

/gene="STY0002"

/note="PS01042 Homoserine dehydrogenase signature"

2801..3730

/gene="STY0003"

/note="synonym: thrB"

2801..3730

/gene="STY0003"

/note="Orthologue of E. coli thrB (KHSE\_ECOLI); Pasta hit to KHSE\_ECOLI (310 aa), 94% identity in 308 aa overlap"

/codon\_start=1

/transl\_table=11

/product="homoserine kinase"

/protein\_id="CAD01156.1"

/db\_xref="GI:16501286"

/db\_xref="GOA:Q8XG95"

/db\_xref="SWISS-PROT:Q8XG95"

/translation="MKVYAPASSANMSVGFVLGAATVPDGTLLGDVVSVAAADHF RLNLGRFADKLPPRENIYOCWERFQALGKTI PVMTLEKNPQIMGLSSACS VVAALVAMNEHCCKPLNDRLALMGEGRISGSIYDNVAFPCPLGGMLMSENGI ISQVFGDEWLWVLAIPGIKVSTAEARAILPAQYRQDCIAGRHLAGIHACYRSRQ POLAAALMKDVIAPRYARLLPGFSQARQAVSEIGALSGISGSGPTLFCDKPETA ORVADMLSKHYLQNSGFGVHCRLDTAGARVVG"

3005..3196

/gene="STY0003"

/note="Pfam match to entry PF00288 GHMP kinases, GHMP kinases putative ATP-binding proteins, score 79.20, E-value 5.7e-21"

3068..3103

/gene="STY0003"

/note="PS00627 GHMP kinases putative ATP-binding domain"

3734..5020

/gene="STY0004"

/note="synonym: thrC"

3734..5020

/gene="STY0004"

/note="Orthologue of E. coli thrC (THRC\_ECOLI); Pasta hit to THRC\_ECOLI (428 aa), 93% identity in 428 aa overlap"

/codon\_start=1

/transl\_table=11

/product="threonine synthase"

/protein\_id="CAD01157.1"

/db\_xref="GI:16501287"

/db\_xref="GOA:Q8Z9R6"

/db\_xref="SPTREMBL:Q8Z9R6"

/translation="MKLYNLKHNEQVSFAQAVTQGLQKQGLFFPHDLPERSLTBID EMNQDFVRSRAKILSAF IGDEIPQILLEVRFAAFAPFAVQVSDVGCLELFFGP TLAFQGGFRVFAQMLTHISGDKPTILTATSGDTGAAVAFYGLNVRVILYPRG KISPLQKELFCITGNIEVADIDDFDAQALVQAFDDEELKTLGLANSINISR LIAQICYFEAVQALPQCARNOIVISVPSGNTGDLTAGLAKSLGLPVRKFTAAVIN DTVPRFLHGKAPKATQATLSNMDVSPNNWPRVEELFRKIKWELTELGAADVDT TQOQWRELKAGYISEPAAVAVARLRQLNPEGEGLFLGTAPAKFKESVESILGE TLALPALAEADPLFLLSHLPADFAALRKLMMTRQ"

3962..4867

/gene="STY0004"

/note="Pfam match to entry PF00291 PALP, Pyridoxal-phosphate dependent enzyme, score 104.50, E-value 2e-27"

4022..4066

/gene="STY0004"

/note="PS00165 Serine/threonine dehydratases pyridoxal-phosphate attachment site" complement(5114..5887)

```

/gene="STV0005"
/notes="synonym: yaaA"
/complement(5114..5887)
/gene="STV0005"
/notes="Orthologue of E. coli yaaA (YAAA_ECOLI); Fasta hit to YAAA_ECOLI (258 aa), 86% identity in 257 aa overlap"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="CAD01158.1"
/db_xref="GI:16501288"
/db_xref="SPTREMBL:Q829R5"
/transl_table="MLTISPATLTDYQSLATRYTOPELLDHSOOLQOARQLSAP
QISLMGISKLADNATREHDPQHPETPDNQAAILAFKGVYVTVGLQAEITADNADP
PAQHLNMLSLGVLRPLDMLQPYRLEMGIRLENPKGLYQFWDIITDKLNEALE
AQQDRVVNLSAEIYFASVPKLDALRIKPFVLDERNGKFKVVSFYAKKARGLMSRF
IIEENLTKPQLTAPDREGYFFDEETSTQDELVFKRYEQ"
/gene="STV0006"
/notes="synonym: yaaJ"
/complement(5966..7396)
/gene="STV0006"
/notes="synonym: yaaJ"
/complement(5966..7396)
/gene="STV0006"
/notes="Similar to Bacillus subtilis amino acid carrier protein aist ALST SW:ALST_BACSU (Q45068; P40743) fasta scores: E(): 0.45.3% id in 477 aa, and to Alteromonas haloplanktis na(+)-linked D-alanine glycine permease DAGA SW:DAGA_ALPHA (P30144) fasta scores: E(): 0.36.5% id in 315 aa
Orthologue of E. coli yaaJ (YAAJ_ECOLI); Fasta hit to YAAJ_ECOLI (476 aa), 76% identity in 476 aa overlap"
/codon_start=1
/transl_table=11
/product="putative amino-acid transport protein"
/protein_id="CAD01159.1"
/db_xref="GI:16501289"
/db_xref="GOA:Q829R4"
/db_xref="SPTREMBL:Q829R4"
/transl_table="MPSEFFSEINILMGSMVYLLGAGCWFWRTGFIQFVRIQFIS
RSLKGSLSQPGGITSQALCTSAAKRGSLNAGVALATAAGPGGAVFWVWSALIG
NATLFAECSLAQLKEDPTQCFRGPANWARGLNWVGFALFLVAYGLIENS
VQANVSRLAHFAFNIPPLISGIALAFCALLIIRGKGVARLMQMLIPLIALIIVAG"

Query Match      47.8%; Score 518.2; DB 1; Length 251050;
Best Local Similarity 67.4%; Pred. No. 1.3e-58; Indels 0; Gaps 0;
Matches 730; Conservative 0; Mismatches 353;

QY 1 ATGCTCTGCTGCTGCGCGAATACCTGCAACAGTTCTACAAAGGCTTGGCGTCTTCCAG 60
Db 147729 ATGTTAGTTTGGCTGCGCGAGCATTTGTCACAAATATTATTCGGCTTTAAAGTCTTTCT 147788

QY 61 TACCTGACCTGCGGGGATCTTCAAGCTGCTCACCGCGCTGCTGCTGCTGCTGCTG 120
Db 147789 TATCTGACGTTTTCGGCGCATGCTGAGCTGTTGACCGCGCTGTCATCTTTTATGATG 147848

QY 121 GGGCCCTGGATGATCGGTACCTTGCAGATGCCCCAGATCGCGCGCGCTGCGCAACGAC 180
Db 147849 GACCCCGATGATGCTGCTGCTGCTGCAAACTCTTTTGGCCAGGTGATGTAACGAT 147908

QY 181 GTTCGCGAGTCCGACTGTCGAAGAGGCGACCCGACATGCGCGCGCGCTGATCTT 240
Db 147909 GGGCCGGAATCGCACTTCAGTAAACCGCGATGCGCGCGATGCGCGCATGATCTG 147968

QY 241 ACGGCCATAGCATGACGACGCTGCTGTGGCGGATCTTCAACCGCTAGCTGTGGTA 300
Db 147969 ACGCGATTGATTTCCGTTCTGTTATGGGTTTACCCTGTTAACCCTAGCTGTGGTC 148028

QY 301 GTGCTGCTGATACCGCTGCTGTCGCTGCCATCGCTGGGTAGACGATACCGAAGTG 360
Db 148029 GTGCTGCTGATTTGATCGCTGACGCGCATTAATCGGTTTGTGATGATGATCGCAAGTG 148088

QY 361 ATCGGAGAAGACTCCCGTGGCTGCGGCGCGCTGGAAGTACTTCTGCGAGTCCGCTGTC 420
Db 148089 GTGCGGAAAGATACCAAGGGCTGATGCGCGCTGGAATATTTCTGGATGCTCGGTTATC 148148

```

```

421 GCATCGCGCGCGCGCTTCTCTACATGACTGCCGAAACCCGATCGAGACCACTG 480
148149 GCGCTCGCGTGGCGCTTTATCTCTGCGGAAAGACACGCCCGCAGCACTG 148208
481 ATCGTGGCGATGCTGAAGAGGTCGAGATCCAGTTGGGCATCTTCTTCTGCTGCTGAC 540
148209 GTGGTGGCGTCTTTAAAGATGTTATGCCGCAATTGGGGCTGTTTACATCTGCTGTC 148268
541 TACTTCGTCATGCTCGGCTCGAGCAATCGGTGAACTCAGGAGGCTCGAGCGCTG 600
148269 TACTTTGTCATGCTCGGTACGGGTAAACGCCGTAACCTGACCGAGGCTTGAATGCTG 148328
601 GCGATCATGCCGACGTAATGTTGCCGCGCGCTGGGCATCTTCTGCTACCTGTCGGCG 660
148329 GCGATTATGCCGACTGCTCTGTTGCCGCGGCTTTCGCTGGGCTTGGCGGACCGG 148388
661 AACGTGAAGTTCGCGAGTACCTGCTGATTCGCAACGTACCGGGCGCGCGAGCTGATC 720
148389 AACATGAATTCGCCAATTACCTGCTATTCGCTATTTACGCCATGCGCGGCGAGCTG 148448
721 GTGTTCTGCGCGCGCTGCTGCGCGCGCTCGGCTTCTCTGCTTCAACACCTATCCG 780
148449 ATTGCTGTACGCGGATTTGTCGCGCGGATTTAGGATTTCTTGTGTTTAAACCTATCCG 148508
781 GCGCAGGTTTTCATGCGGCAAGCTGCGCGCTGCGCTGGCGCGCGCTGGCGGACCAATC 840
148509 GCGCAGGTTTTCATGCGGCAAGCTGCGGCTGCGCTGGCGCGCGCTGGCGGACCAATC 148568
841 GCGGTGATGCTGCGCGGACGAGATCGTGTGTTTCATCATGGGTGGGCTGTTCTCATGAA 900
148569 GCGGTGCTGCTGCGTACGGAATTTCTGCTGTGATCATGGGCGGCTCTTTGTGGTGA 148628
901 ACCCTCTCGTGTGATGATCCAGTCTGCTTCTTCAAGCTGACGCGCGCGCTTCTCGT 960
148629 ACTCTGCTGCTGATCTCTGCGGCTGCTTCTTAACTACGCGGACGAGCTATTTTCGT 148688
961 ATGGCGCCCATCATCACCATTTCGAATGAAAGCTGCGCGGACCGCGCGCTGATCGT 1020
148689 ATGGCGCCCATCATCACCATTTCGAATGAAAGCTGCGCGGACCGCGCGCTGATCGT 148748
1021 CGTTCTGATCATCACCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1080
148749 CGTTCTGATCATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 148808
1081 TGA 1083
148809 TAA 148811

RESULT 10
AB016834
LOCUS
DEFINITION
Salmonella enterica subsp. enterica serovar Typhi Ty2, section 1 of 16 of the complete genome.
ACCESSION
AB016834 AB014613
VERSION
AB016834.1 GI:29136167
KEYWORDS
Salmonella enterica subsp. enterica serovar Typhi Ty2
SOURCE
Salmonella enterica subsp. enterica serovar Typhi Ty2
ORGANISM
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales; Enterobacteriaceae; Salmonella.
REFERENCE
1 (bases 1 to 300169)
Deng W., Liou S.-R., Plunkett III, G., Mayhew, G.F., Rose, D.J., Burland, V., Kodoyianni, V., Schwartz, D.C. and Blattner, F.R., Comparative Genomics of Salmonella enterica Serovar Typhi Strains Ty2 and CT18
J. Bacteriol. 185 (7), 2330-2337 (2003)
MEDLINE
22531367
PUBMED
12644504
REFERENCE
2 (bases 1 to 300169)
Deng W., Liou S.-R., Plunkett, G. III, Mayhew, G.F., Rose, D.J., Burland, V., Kodoyianni, V., Schwartz, D.C. and Blattner, F.R.

```

TITLE Direct Submission  
JOURNAL Submitted (25-SEP-2002) Laboratory of Genetics, University of Wisconsin - Madison, 445 Henry Mall, Madison, WI 53706, USA  
FEATURES  
Location/Qualifiers  
1. .300169  
/organism="Salmonella enterica subsp. enterica serovar Typhi Ty2"  
/mol\_type="genomic DNA"  
/strain="Ty2"  
/serovar="Typhi"  
/sub\_species="enterica"  
/db\_xref="taxon:209261"  
190. .255  
/genes="thrL"  
/locus tag="t0001"  
190. .255  
/genes="thrL"  
/locus tag="t0001"  
/notes="corresponds to STY0001 from Accession AL513382: Salmonella typhi Ctr18"  
/codon\_start=1  
/transl\_table=11  
/product="thr operon leader peptide"  
/protein\_id="AAO67735.1"  
/db\_xref="GI:29136168"  
/translation="MNRISTTTTITTTTGGAG"  
337. .2799  
/genes="thrA"  
/locus tag="t0002"  
337. .2799  
/genes="thrA"  
/locus tag="t0002"  
/notes="corresponds to STY0002 from Accession AL513382: Salmonella typhi Ctr18; homoserine dehydrogenase I"  
/codon\_start=1  
/transl\_table=11  
/product="aspartokinase I"  
/protein\_id="AAO67736.1"  
/db\_xref="GI:29136169"  
/translation="MRVLFGTTSVANAERFLRVADILSNRQGVATVLSAPAKIT NHLVMIETKIGQDAPNISDAERIFSDLLAGASAQFGFLRMKVMVEFFAQIK HVLHGISLLQCDSINAAICRGKNSAIWAGLEARGHVTVIDPVKLLAVGHY LESTVDIAESTRIASQIPADHMIWMAGTAGNEKGEVLVGRNGSDYSAVLAACL RADCCILWTDVYVTCDPQVDPDARLLKMSYQEMLSYFGAKVLHPRTTIPAQF QICLLIKVGNPQAPGLTGASDDNLPKGISLNINWAMFSVGFPMKMGIMGAAR VFAMRAGISVVLITQSSSEYSISFCVQSDCARRAMQDEFYLEKEGLEFLAV TERLAIISVVDGDMRTLRIGISAKFFAALARANINVAIAQSSERSISVYVNNDDATT GVRVTHQMLFNTDQVIEVFVIGVGVGALLQKQOQTLKXKHIDLRVCGVANSKA LITNVHGLDNWQAIAQANAPENGLRLVKEVHLLNPVIVCTSSQAVADQYAD FLREGFHVTPNKKANTSMNDYVHOLREAAQSRKFLVDTNVGAGLPVLENQLLN AGDELQKSGILSGLSFFIFGLEKGSLSQATALARENGYTEPDPDRDLSDMVARK LLLIARETGRELSLDIVIEPVLDFDASGDVTAFAHLPQLDDAFAARVAKARDEG KLVRYGVNIEEDGVCVKIAEVDGNDPLFKVKNXGLENALAFYSHYYQPLVLVRGYGAG NDVTAAGVFADLLRTLSWKLGV"  
2801. .3730  
/genes="thrB"  
/locus tag="t0003"  
2801. .3730  
/genes="thrB"  
/locus tag="t0003"  
/notes="corresponds to STY0003 from Accession AL513382: Salmonella typhi Ctr18"  
/codon\_start=1  
/transl\_table=11  
/product="homoserine kinase"  
/protein\_id="AAO67737.1"  
/db\_xref="GI:29136170"  
/translation="MVKVAPASSANNMVGFDVLAGAVTPVDGTLLGDVVSVEAADHF RLNLGRFADLPPEPENVIVOCWERFQCALQTPVAMTEKMPISGSLGSSACS VVAALVAMNEHCCKPLNDTLALMGELESGRISGTHYDNVAPCTLGGWQLMEINGI ISQVPCGFDEWLVAYPGIKVSTABARAILPAQVRODCIAGRHLAGFIHACYSRQ PQALAAKMDVIAEPFRALLPGFSQARQAVSEIGALASGISGSGFTLFCCKPETA QRVADWLSKHYLQNEQGFHICKLDTAGARVVG"

gene  
CDS  
gene  
CDS  
gene  
CDS  
gene  
CDS

3734. .5020  
/gene="thrC"  
/locus tag="t0004"  
3734. .5020  
/gene="thrC"  
/locus tag="t0004"  
/notes="corresponds to STY0004 from Accession AL513382: Salmonella typhi Ctr18"  
/codon\_start=1  
/transl\_table=11  
/product="threonine synthase"  
/protein\_id="AAO67738.1"  
/db\_xref="GI:29136171"  
/translation="MKLYNLKDNHNEOVSEFAQVNTQGLGKQGLFPFPHDLPEFSLTEID EMLNQDVSSSAKILSAFIDGRIEPOQILERYRAAFAPAPVAQVSDVGCLELPHGP TIAFDGGRFMAQMLHISGDKREVITLITATSGDTGAFAFAFGLGNRVVILYPRG KISPLQKLFCTIGGNIETVAIDGDFDACALVKQAFDDEELKATGLGNSANINIR LLAQICYFEAVAPQAGARNQLVISVSGNFGDLTAGLLAKSLGLVKRPIAAATNIN DTDFRFLDGGWAKPAQTATLSNAMDVSQNNWPRVEELFRKIRWLTGCGYAAVDOT TTQOTMRELKAGKVISSPHAAVAYRALRDOLNPGCEYGLGLGTAHPAKFKESVESILGE TLALPEALASRADILPLLSHLPADPAALRKLMMTRQ"  
complement(5114. .5887)  
/gene="yaaA"  
/locus tag="t0005"  
complement(5114. .5887)  
/gene="yaaA"  
/locus tag="t0005"  
/notes="corresponds to STY0005 from Accession AL513382: Salmonella typhi Ctr18"  
/codon\_start=1  
/transl\_table=11  
/product="conserved hypothetical protein"  
/protein\_id="AAO67739.1"  
/db\_xref="GI:29136172"  
/translation="MLILISPAKTLDYQSLPATTTRYTQPELLDHSQOLIOCARLSAP QISRLMGISXKADLNATRFHDWQPHFTPNARQAILAFKGDVYVTGLQAEFNDADPD FAQHLRLMSGLYGLRPLDLQMPYRLMGIRLENPRKGLDIQFWGDIITDKLNEALE AQDGVVYNLASEEYFKSVKPKDDAELIKPFLDENRKGPKVYVSVYAKKARGLMSRF IENRLTKPELTAFDREGYFFDSETSQDELVFRYEQ"  
complement(5966. .7396)  
/gene="yaaY"  
/locus tag="t0006"  
complement(5966. .7396)  
/gene="yaaY"  
/locus tag="t0006"  
/notes="corresponds to STY0006 from Accession AL513382: Salmonella typhi Ctr18"  
/codon\_start=1  
/transl\_table=11  
/product="putative amino-acid transport protein"  
/protein\_id="AAO67740.1"  
/db\_xref="GI:29136173"  
/translation="MPEFFSFINEILWGSVMYILLGAGCWFTRTGTGFIQFRIYRQPS RSLKGSLSPOPGLTSFQALCTSLAARGSNLAGVALAIAAGGPGAVFWMVWGAIIIG MATSFAECSLAQLYKERDPTGQFRGGPWWMGVWVFWALFLVAYGLIFNS VQANAVSRALHFAENIPPLISGIALAFCALLIIRGIKGVARLQMLPLIALLAVAG SVFICLWHIQMPGCVIASIVKSAFQWQBAAGAGYTLTQAITSGFQFGMFSNAGWG STPNAAAASVPHPPVAQGVQMGIVFSDTIITCTASAMILLAGHSHASSTEGIQ LLOHMAVSLTGEWGSFVALIVILFAFSSIVANYIYAENLFFLRHNHAKAILWLRLA TLGHVIAGTLISFPLIWLADMINMACWAINLTALLLSPVVYTAGDYLRQRLGVR PQDFPRFPDIEPQLAPDTWDAASRD"  
7665. .8618  
/gene="talB"  
/locus tag="t0007"  
7665. .8618  
/gene="talB"  
/locus tag="t0007"  
/notes="corresponds to STY0007 from Accession AL513382: Salmonella typhi Ctr18"  
/codon\_start=1  
/transl\_table=11  
/product="transaldolase B"  
/protein\_id="AAO67741.1"

Query Match	47.8%;	Score 518.2;	DB 1;	Length 300169;
Best Local Similarity	67.4%;	Pred. No. 1.3e-59;		
Matches 730;	Conservative 0;	Mismatches 353;	Indels 0;	Gaps 0;
QY 1	ATGCTCCTGCTGCTGCGCGCAATACCTGCAACAGTCTCTACAAGGGCTTCGGCGTCTCCAG 50			
Db 147721	ATGTTAGTTTGGCTGGCGAGCATTTGGTCAAAATTATTCCGGCTTTAAACGCTTTTCTT 147780			
QY 61	TACCTGACCCCTGCGGGCATTCTCAGCGTGTCTACCGCGTGTCCGTGTGCTGTGGCTG 120			
Db 147781	TATCTGACGTTTCGGGCCATCGTCAGCCTGTGTACCGCGCTGTCATCTCTTATGGAATG 147840			
QY 121	GGGCGCTTGATGATCTGCTGATACCTTGAGATATCCCCAGATATGGCCAGGCGGTGCCAACGAC 180			
Db 147841	GGCCGCGGTATGATCGCTCGTCTGCAAAAATCTCTTTTGGCAGGTGGTACGTAACGAT 147900			
QY 181	GGTCGCGAGTCGCACTCTGCAAGAGGGCACCCGACCATATGGCGCGCGCCCTGATCCCTT 240			
Db 147901	GGCCCGGAATCGCATTTGATTAACCGCGTACGCCAGATATGGCGGCGCATCATGATCCTG 147960			
QY 241	ACCGCATAGCCATCAGCACGCTGTGTGGCGGATCTTTCCAAACCGCTACGTTGTGGGTA 300			
Db 147961	ACGGCGATTGTATTTCCGTTCTGTATATGGGCTTACCGCTTAACCGCTACGTTGCTGTC 148020			
QY 301	GTGCTGCTGTTACCCCTGCTGTTTGGTGGCCATATGGCTGGGTAGACGACTACCGCAAGGTG 360			
Db 148021	GTGCTGCTGTTATGATCGGCTACGGCATTTATCGGCTTTTGTGATGATGACTACCGCAAAATG 148080			
QY 361	ATCGAGAAGAACTCCCGTGGGCTGCCGAGCGCTGGGAAGTACTTCTGGCAGTGGGTGTC 420			
Db 148081	GTGCGGAAGAGTACCAAGGGCTGATTCGGCGCTGGAAATATTTCTGGATGTGGGTATC 148140			
QY 421	GGCATCGGCGCGCGTGTTCCTCTACATGACTGCGGAAACCCCGATCGAGACCAACCTG 480			
Db 148141	GGGCTCGGCTGGGCTTTTGGCTTTATCTGCTCGGAAGAGACACGCCCGGACCCAACTG 148200			
QY 481	ATCGTCCGATGCTGAAGAGCGTCGAGATATCAGTTGGGCATCTTCTTGTGGTCTGTAAC 540			
Db 148201	GTGGTCCGCTCTCTTTAAAGATGTTATGCCGAATGGGGCTGTTTACATCTTGTGCTGCC 148260			

VVPFFKDVMPQLPYLLAYFYVIVGTGNVNLTDGLDGLAIMPTVFVAGGFALVAVM  
TGNMFASVYHLPYLRHAGELVIVCTAIAGAGLGFWENTYPAQVFMGDVGSALGGA  
LGLIALLRQEFLLVIMGVFVETLSVILQVGSFKLRQRIFRAPAPIHHHYELKWP  
EPRVIVFWIISLMVLIGLATLKR"

ORIGIN

Query Match 46.4%; Score 502.2; DB 6; Length 1083;  
Best Local Similarity 66.5%; Pred. No. 4.1e-56;  
Matches 720; Conservative 0; Mismatches 363; Indels 0; Gaps 0;

QY 1 ATGCTCCTGCTGTCGCGGATACCTGCAACAGTCTACAGGGCTTCGGCTCTTCAG 60  
DB 1 ATGTTAGTTTGGCTGGCGGACATTTGGTCAAAATATATTCCGGCTTTAACGTCTTTCC 60

QY 61 TACCTGACCTGCGCGGCAATTTCTCAGCGTGTCTCAGCGCTGTGCGTGTGCGTGGTG 120  
DB 61 TATCTGACGTTTCGCGCCATCGTCAGCGCTGTGACCGCTGTTCATCTCATTTGGATG 120

QY 121 GGGCCCTGATGATCGTACCTTGCAGATCCCGCAGATCGGCAGCGCTGCGCAACGAC 180  
DB 121 GGGCCCGGTATGATGCTCATTTTGCAAAACCTTTCTTTGGTCAGGTGGTGAACGAC 180

QY 181 GGTCCCGAGTCGACCTGTGCGAAGAGGCGACCCCGACCATGCGCGGCCCTGATCCTT 240  
DB 181 GGTCTGATACACTTTCAGCAAGCGGTGACCGACCATGGCGGGATATGATCTCTG 240

QY 241 ACGGCATAGCCATAGCAGCGTGTGTGGCGGATCTTTCCAAACGCTACGTGGGTA 300  
DB 241 ACGGCATTTGATCTCCGCTACTGCTGTGGGTGTACCCGTCCAATCCGTACGTCTGGTGC 300

QY 301 GTGCTGCTGTACCTGCTGTGCTGCTGCATCGCTGGGTAGAGACTACCGCAAGTG 360  
DB 301 GTGTTGGTGTGTGTAGTTACGTGATGTTATGTTGTTGTTGATGATATTCGCAAGTG 360

QY 361 ATGAGAAGAACTCCGTCGCTCGCGAGCGCGTGGAGTACTTCTGGCAGTCGCTGTC 420  
DB 361 GTGCGTAAAGACACCAAGGTTGATCGCTGTGTGAAGTATTTCTGATGTGCTCAT 420

QY 421 GGCATCGCGCGCGCTGTCTCTACATGATCGCGAACCCTGATCGAGACCCCTG 480  
DB 421 GGCCTGGGTGTGCGCTGTGCGCTGTACCTTGGCGCAAGACACGCCCGCAACGCGAGTG 480

QY 481 ATGCTGCCGATGCTGAAGACGCTCGAGATCCAGTTGGGCACTTCTTCTGCTGCTGACC 540  
DB 481 GTGGTCCCATTTCTTTAAGATGTGATGCGCGAGCTGGGCGTGTCTACATTTCTGCTGCT 540

QY 541 TACTTGTGATGTCGCTGCGAGCAATGCGGTGAACCTCAGCAGGTCTCAGCGCCTG 600  
DB 541 TACTTGTGATGTTGGGTACTGCGAACGCGGTAAACCTGACCGATGCTCAGCGCCTG 600

QY 601 GCGATCATGCGGACGCTGAATGTTGCGCGCGCTGGGCACTTCTGCTACCTGTGCGGC 660  
DB 601 GCAATTATGCGGACCGTATTTGTCGCGGTGTTGCGCTGGTGGCGTGGCGGACCGGC 660

QY 661 AAGTGAAGTTTCCGAGTACTGCTGATATCCGATATCTCGGACACGCCGGGAACTGGT 720  
DB 661 AATATGAATTTTCCGAGTACTTGCATATACCGTATCTCGGACACGCCGGGAACTGGT 720

QY 721 GTGTTCTGCGCGCTGTGTCGCGCGCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780  
DB 721 ATTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780

QY 781 GCGCAGTCTTTCATGCGGACGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCT 840  
DB 781 GCGCAGTCTTTCATGCGGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840

QY 841 GCGGTGATGTCGCGGACGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900  
DB 841 GCGGTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900

QY 901 ACCCTCTGCTGATGATCAGGTCGCTTCTTCAAGCTGACCGGACCGCGCTGCTTCCGT 960  
DB 901 ACGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960

961 ATGCGCGCATCATCACCATTTCGAACTGAAAGCTGGCGGACCGCGCGTGCATCGT 1020  
DB 961 ATGCAACGATTCATCACCATTTCGAACTGAAAGCTGGCGGACCGCGCGTGCATCGT 1020

QY 1021 CGTTCTGATCATCACCATTTCGATGCTGCTGATCGCGCTGCGCACCTTGAAGCTCGT 1080  
DB 1021 CGTTCTGATCATCACCATTTCGATGCTGCTGATGCTGCTGCGCACCTTGAAGCTCGT 1080

QY 1081 TGA 1083  
DB 1081 TAA 1083

RESULT 12  
ECMUROY  
LOCUS Escherichia coli murd gene and ORF-Y (EC 6.3.2.9).  
DEFINITION Escherichia coli murd gene and ORF-Y (EC 6.3.2.9).  
ACCESSION X51584.X5117  
VERSION X51584.1 GI:42058  
KEYWORDS membrane protein; murd gene;  
UDP-N-acetyl-muramoyl-L-alanyl-D-glutamate synthetase; unidentified  
reading frame.  
SOURCE Escherichia coli  
ORGANISM Escherichia coli  
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
Enterobacteriaceae; Escherichia.  
REFERENCE 1 (bases 1 to 2608)  
AUTHORS Ikeda,M., Wachi,M., Ishino,F. and Matsuhashi,M.  
TITLE Nucleotide sequence involving murd and an open reading frame ORF-Y  
spacing murf and ftsW in Escherichia coli  
JOURNAL Nucleic Acids Res. 18 (4), 1058 (1990)  
MEDLINE 90192099  
PUBMED 2179861  
REFERENCE 2 (bases 1 to 2608)  
AUTHORS Ikeda,M.  
TITLE Direct Submission  
JOURNAL Submitted (02-JAN-1989) Ikeda M., Institute of Applied  
Microbiology, The University of Tokyo, 1-1-1 Yayoi, Bunkyo-ku,  
Tokyo 113, Japan  
REFERENCE 3 (bases 26 to 1354)  
AUTHORS Flouret,B.  
TITLE Direct Submission  
JOURNAL Submitted (15-MAR-1990) Flouret B., URA 1131 du Centre National de  
la Recherche Scientifique, C.N.R.S., Biochimie Moléculaire et  
Cellulaire, Batiment 432 Université Paris-Sud, 91405 Orsay, France  
COMMENT Sequence independently determined by [3], but authors accept  
sequence [1] as definitive.  
FEATURES  
source Location/Qualifiers  
1..2608 /organism="Escherichia coli"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:562"  
/map="2min"  
98..102 /note="pot. ribosome binding site"  
111..1193 /note="unnamed protein product; ORF-Y (AA 1-360)"  
/codon\_start=1  
/transl\_table=11  
/protein\_id="CAA35932.1"  
/db\_xref="GI:42059"  
/db\_xref="GOA:P15876"  
/db\_xref="SWISS-PROT:P15876"  
/translation="MLVLAELHVKYSGNFVSYLTFRATVLSLLTALFTSLWMGPRM  
IAHLQKLSFGQVNRNDGSPHESKRGPTMGIMILTAIVISVLLWAYPSNPVYCVL  
VVLVGVIGFVDDYKVKVRKTKGLIARKYFWMVSIVIALGVAFYLAGKDTIPATQL  
VVPFFKDVMPQLGFLYLLAYFYVIVGTGNVNLTDGLDGLAIMPTVFVAGGFALVAVM  
TGNMFASVYHLPYLRHAGELVIVCTAIAGAGLGFWENTYPAQVFMGDVGSALGGA  
LGLIALLRQEFLLVIMGVFVETLSVILQVGSFKLRQRIFRAPAPIHHHYELKWP  
EPRVIVFWIISLMVLIGLATLKR"  
1018..1023  
/note="pot. -35 region"

promoter

```

promoter 1043..1048 /note="pot. -10 region"
RBS 1181..1185 /note="pot. ribosome binding site"
CDS 1196..2512 /note="unnamed protein product; murD gene product (AA
1-438)"
/codon_start=1
/transl_table=11
/protein_id="CAA35933.1"
/db_xref="GI:42060"
/db_xref="GOA:P14900"
/db_xref="SWISS-PROT:P14900"
/translation="MADYQGVNVIIGLITGLSCVDFFLARGVTPRVMDTRMTPPGL
DKLEAVERHTGSNDLWMAADLIVASPGIALAHPSLSAADAGIIEVDIELFQRE
ACAPVATGNSGKSTVTLVGEWAAGVNVGNGIPLALMLIDDECELVLELS
SFQLETSLSQAAVATILNVTEDHMDRYPRGLQYRAKLRIYENAKVCVNRADALT
MPIGABERCVSFVNMWDHNLHQQGETWLVRGKVLNVKEMKSGQNYNTALAA
LALADAGLPASIKALITFTGLPHRFVLEHNGVRWINDSKATNVGSTEALNGL
HVDGTHLLGGDDKSDAFPLARYLNGDNVRLYCFGRDGAQLAALRPEVAQETME
QAMRLAPRVQPGDMVLLSPACASLDQFNKFNFORGNFARLAKELG"

ORIGIN
Query Match 46.4%; Score 502.2; DB 1; Length 2608;
Best Local Similarity 66.5%; Pred. No. 3.5e-56;
Matches 720; Conservative 0; Mismatches 363; Indels 0; Gaps 0;
QY 1 ATGCTCTGCTGCTGGCGGAATACCTGCAACAGTCTTACAAGGCTTCGGCGTCTCCAG 60
DB 111 ATGTTAGTTTGGCTGGCGGAACATTTGTCAAATATTATTCCGGCTTTAAGCTTTTCC 170
QY 61 TACCTGACCTTGGCGGGAATCTCAGGTGCTCACCGGCTGTCTGCTGTGCTGTGCTG 120
DB 171 TATCTGAGTTTGGCGGCAATCGTCAGCGCTGTGACCGGCTGTTCATCTCATTTGGATG 230
QY 121 GGGCGCTGATGATCCGTCACCTTGCAGATCCCCAGATCGGCCAGCGCGTGGCAACGAC 180
DB 231 GGGCGGATGATGATGCTCATTTGCAAAAATCTTCTTGGTCAAGTGGTGGTAAACGAC 290
QY 181 GGTGCGAGTCGACGCTGCAAGAGGCGACCCGACCATGAGCGCGGCGGCGCTGATCCTT 240
DB 291 GGTCTGAATCACACTTCAGCAAGCGGATGACCGGCGACCATGGCGGGGATTATGATCCTG 350
QY 241 ACCGCAATAGCATCAGCACGCTGCTGGCGGATCTTCCACCGCTACGTTGGTGTA 300
DB 351 ACGCGATTGATCTCGTACTGCTGTGGCTTACCGGCTCAATCGCTAGCTGTGCTG 410
QY 301 GTGCTGCTGATACCTGCTGCTCGGTGCCATCGGCTGGGTAGACGACTACCGCAAGGTG 360
DB 411 GTGTTGGTGGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 470
QY 361 ATCGAGAGAACTCCCGTGGCTCCGAGCGCTGGAGTACTTCTGCGAGTGGTGTTC 420
DB 471 GTGGTAAAGACACCAAGGTTGATCGCTGTTGGAGTATTTCTGGATGCGGTCAAT 530
QY 421 GGCATCGCGCGCGGCTGCTCTCTACATGACTCCGAAACCCGATCGAGACACACCTG 480
DB 531 GCGCTGGGTGCGCTTCCGCTGTACCTTCCCGGCAAGACACGCCCGCAACGACGCTG 590
QY 481 ATCTGCGCGATCGTCAAGAGGCTCGAGATCAGTGGGCTCTTCTTCTGCTGCTGACC 540
DB 591 GTGTGCCCATTTTAAAGATGTGATGCCGAGCTGGGGCTGTCTACATCTGCTGGCT 650
QY 541 TACTTCTGCTATCGTGGCTCGAGCAATCGGTGAACCTCACCGACGGTCTCGAGCGGCTG 600
DB 651 TACTTCTGCTATTTGGGTACTGGCAACCGGTAACCTGACCGATGTCTCGAGCGGCTG 710
QY 601 GCGATCAGCGAGGTAAATGGTTCGCGCGCGCTGGCGATCTTCTGCTACCTGTGCGG 660
DB 711 GCAATTAAGCGGACCGTATTTGTGCGCGGTGTGTTTGGCTGTGGGTGGGACCGGC 770
QY 661 AACGTGAAGTTCCGCGAGTACTGCTGATTCCTCAACAGTACCGGCGCGCGAGCTGATC 720

```

```

DB 771 AATATGAACCTTTGCCAGCTACTTGCATATACCGTATCTGGACACGCCGGGAACCTGTT 830
QY 721 GTGTTCTGGCGCGCTGGTGGCGCGCTCGGCTTCTCTCTGTTCAACACCTATCCG 780
DB 831 ATTGTCTGTACCGCATAGTACGGGCGAGGACTGGGCTTCTCTGTTTAAACACCTATCCG 890
QY 781 GGCAGGCTTTCATGGCGAGCTGGCGCGCTGGCGCTGGGCGCGCGCTGGGCGGCACTATC 840
DB 891 GGCAGGCTTTCATGGCGAGTGTAGTTCGCTGGCGCTGGGCTGGGCTGGGCTGGGCTATTC 950
QY 841 GCGGTGATCGTGGCGGAGAGATCGTGTGTTTCATCGGTGGGCTGGGCTGGGCTGGGCTATTC 900
DB 951 GCGGTACTGCTACGTGAGGATTCCTGCTGGTGAATTATGGGCGGCTGTTCTGCTGATGAA 1010
QY 901 ACCCTCTGCTGATCATCCAGTCCGCTTCTTCAAGCTGACCGGACCGCGGCTCTCCGT 960
DB 1011 AGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTTAACTGCGGCGGCAAGTATTTCCGC 1070
QY 961 ATGGCGCGGATCCATCACCATTTCGATTCGAAAGCTGGCGGACCGCGGCTGATCGTG 1020
DB 1071 ATGGCACCGGATTCATCACCCTATGAACTGAAAGCTGGCGGAAACCGCGGCTCATTGTG 1130
QY 1021 CGCTTCTGATCATCACCCTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1080
DB 1131 CGCTTCTGATTCATTCGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1190
QY 1081 TGA 1083
DB 1191 TAA 1193

RESULT 13
AE000118
LOCUS
DEFINITION
Escherichia coli K12 MG1655 section 8 of 400 of the complete
genome.
ACCESSION
AE000118 U00096
VERSION
AE000118.1 GI:1786262
KEYWORDS
Escherichia coli K12
Escherichia coli K12
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
Enterobacteriaceae; Escherichia.
REFERENCE
1 (bases 1 to 21757)
Blattner, F.R.; Plunkett, G. III, Bloch, C.A., Perna, N.T., Burland, V.,
Kiley, M., Collado-Vides, J., Glasner, J.D., Rode, C.K., Mayhew, G.F.,
Gregor, J., Davis, N.W., Kirkpatrick, H.A., Goeden, M.A., Rose, D.J.,
Mau, B. and Shao, Y.
The complete genome sequence of Escherichia coli K-12
Science 277 (5331), 1453-1474 (1997)
97426617
MEDLINE
2 (bases 1 to 21757)
9278503
PUBMED
REFERENCE
2 (bases 1 to 21757)
Blattner, F.R.
Direct Submission
Submitted (16-JAN-1997) Guy Plunkett III, Laboratory of Genetics,
University of Wisconsin, 445 Henry Mall, Madison, WI 53706, USA.
Email: ecoligenetics.wisc.edu Phone: 608-262-2534 Fax:
608-262-7459
3 (bases 1 to 21757)
Blattner, F.R.
Direct Submission
Submitted (02-SEP-1997) Guy Plunkett III, Laboratory of Genetics,
University of Wisconsin, 445 Henry Mall, Madison, WI 53706, USA.
Email: ecoligenetics.wisc.edu Phone: 608-262-2534 Fax:
608-262-7459
4 (bases 1 to 21757)
Plunkett, G. III.
Direct Submission
Submitted (13-OCT-1998) Laboratory of Genetics, University of
Wisconsin, 445 Henry Mall, Madison, WI 53706, USA
This sequence was determined by the E. coli Genome Project at the
University of Wisconsin-Madison (Frederick R. Blattner, director).

```

Supported by NIH grants HG00301 and HG01428 (from the Human Genome Project and NCHGR). The entire sequence was independently determined from E. coli K12 strain MG1655. Predicted open reading frames were determined using GeneMark software, kindly supplied by Mark Borodovsky, Georgia Institute of Technology, Atlanta, GA, 30332 (e-mail: mark@embarcadero.gatech.edu). Open reading frames that have been correlated with genetic loci are being annotated with CG Site Nos., unique ID nos. for the genes in the E. coli Genetic Stock Center (CGSC) database at Yale University, kindly supplied by Mary Berlyn. A public version of the database is accessible (<http://cgsc.biology.yale.edu>). Annotation of the genome is an ongoing task whose goal is to make the genome sequence more useful by correlating it with other data. Comments to the authors are appreciated. Updated information will be available at the E. coli Genome Project's World Wide Web site (<http://www.genetics.wisc.edu>). \*\*\* The 3. coli K12 sequence and its annotations are periodically updated; this is version M54. No sequence changes. Annotation updates: updated gene identifications and products; all new functional assignments courtesy of Monica Riley; added promoters, protein binding sites, and repeated sequences described in reference 1; the unique numeric identifiers beginning with a lowercase 'b' assigned to each gene (protein- or RNA-encoding) are now designated as gene synonyms instead of labels. This should allow them to be searched for in Entrez as gene names.

## FEATURES

Source	Location/Qualifiers
gene	1..21757 /organism="Escherichia coli K12" /mol_type="genomic DNA" /strain="K12" /sub_strain="MG1655" /db_xref="taxon:83333" complement(77..163) /gene="leuL" /notes="synonym: b0075" complement(77..163) /gene="leuL" /function="leader; Amino acid biosynthesis: Leucine" /notes="f28; 100 pct identical to LPL_ECOLI SW: P09149" /codon_start=1 /transl_table=11 /product="leu operon leader peptide" /protein_id="AAC73186.1" /db_xref="GI:1786263" /translation="MTHIVRFIQLLLNASSLRGRVSGIQH" complement(197..226) /notes="factor Sigma70; promoter leu; documented +1 at83735"
protein_bind	complement(201..221) /notes="No predicted promoter" /bound_moiety="LexA predicted site"
protein_bind	complement(201..221) /notes="central position to predicted promoter: -0.5"
protein_bind	complement(354..368) /notes="No predicted promoter"
protein_bind	complement(354..368) /bound_moiety="Lrp documented site"
protein_bind	complement(354..368) /notes="central position to leuABCD promoter: -156.5"
protein_bind	460..474 /bound_moiety="Lrp documented site"
promoter	/notes="central position to predicted promoter: -28.5" /bound_moiety="Lrp predicted site"
protein_bind	462..490 /notes="factor Sigma70; predicted +1 start at 84042" complement(519..537) /notes="central position to predicted promoter: -319.5"
protein_bind	554..565 /bound_moiety="tyrR predicted site"
gene	/notes="central position to predicted promoter: 64" /bound_moiety="AraC predicted site" 646..1767 /gene="leuO"
CDS	/notes="synonym: b0076" 646..1767 /gene="leuO" /function="putative regulator; Amino acid biosynthesis: Leucine" /notes="o373; 100 pct identical to LEUO_ECOLI SW: P10151 (290 aa) but contains 59 additional N-ter aa and 24 C-term residues; 100 pct identical to PIR: S40589 but contains 24 additional C-term residues" /codon_start=1 /transl_table=11 /product="probable transcriptional activator for leuABCD operon" /protein_id="AAC73187.1" /db_xref="GI:1786264" /translation="MTHSTAMDSVFTRTFMFSEYSCFFFLFMHDKSYSSGLFLFCIPREBELSVTVELSPVQNDPETAELSKPQRMVDNLNLLTFPVAWQEQNITRAAHVLGMQPAVSNAVARLKVNFNDLFPVYRGKIQPTARAFQPLFGSVRAQLQVONELPQGSEFASERVVFLHVCSPDLDSILTSQIVNHIEQIAPNHVMFKSLNONTBOLRYQETFEVISEDHFRPSTVPLPKDEMVLVASKNHPTIKGPLAKHVDYNEQHAASLDRFASFQGPWYDITDKASIAIYQGMAMSVLSVVSQTHLVALAPRWLAEPFAESLELQVLPPLKQNSRTCYLSWHEAAGRDKGHWMEEQVLSICKR" /notes="central position to predicted promoter: -39.5" /bound_moiety="Lrp predicted site"
promoter	1813..1839 /notes="factor Sigma70; promoter ilvIHp4; documented+1 at 85394"
protein_bind	1833..1847 /notes="central position to predicted promoter: -8.5" /bound_moiety="Lrp predicted site"
promoter	1838..1866 /notes="factor Sigma70; promoter ilvIHp3; documented+1 at 85420"
protein_bind	1915..1929 /notes="central position to predicted promoter: 73.5" /bound_moiety="Lrp predicted site"
protein_bind	1949..1963 /notes="central position to predicted promoter: 107.5" /bound_moiety="Lrp predicted site"
promoter	1955..1984 /notes="factor Sigma70; promoter ilvIHp2; documented+1 at 85534"
protein_bind	1978..1992 /notes="central position to predicted promoter: 136.5" /bound_moiety="Lrp predicted site"
gene	1995..3809 /gene="ilvI" /notes="synonym: b0077" 1995..3809 /gene="ilvI" /EC_number="4.1.3.18" /function="enzyme; Amino acid biosynthesis: Isoleucine, Valine"
CDS	/notes="o604; 98 pct identical (1 gap) to 522 residues from ilvI_ECOLI SW: P00893 (566 aa)" from ilvI_ECOLI SW: P00893 (566 aa)" /codon_start=1 /transl_table=11 /product="acetolactate synthase III, valine sensitive, large subunit" /protein_id="AAC73188.1" /db_xref="GI:1786265" /translation="MFMSLAANCLSKIGRLMCTHPFQTVQAMMLSCAMVVRSLIDQGVQVFGPGGAVLDIYDALHTVGGIDHVLVRHEQAARVHADGLARATGTVGVVLVTSGFGATNAITGIATYMDSIPLVLSGVATSLIGYDAFQECMDVGI8RPVVKHSF LVTGTEIPQVLKAFWLAASGPGVVDLPKDILNPANKLPYWPESVMSRSHPTTKTHGQIKRALOTLVAAKPVPVVGGAITAGCHQKLTETBALNLPVVCSLMGLGAPPAATHRCALGMGMHGTVEANMTMHNADVI FAVGVRFDRTTNNLAKVCNATLHID IDPTSI8KTVTADIPVGDARQVLEOMLELISOESAHQPLDEIDMWQQL8QEAROC LKDYTHSEKI.KPQAVIETLWRLTKGDYVTSVSDGQHQMPAALYFPFKPRRNINSGSL GTMGFGPALGMPALPEETVVCVTGDSIQMNIQELSTALQLEPVLVNLNLRRL GMVQWQMDIYSGRHSQSLSLDFVRLAEAYGHVGIQISHPELESKLSLEAL8QVR NNLRFVDVTVDSERHVPYMQIRGGGMDENWLSKTERT"

Db	13237	GGCAGGCTTTATGGGCGATGATAGGTTGCTGGCGGTTAGTGTGGTTCAGTATATC	13299
Qy	841	GGGGTGATCGTGGCCAGGAGATCGTGTCTTTCATCATGGGTGGGTGTTCTGTCATGAA	900
Db	13297	GGCGTACTCTACGTACGAGATTCCTGTCTGGTGAATTATGGGGGGCGTGTTCGTGGTAGAA	13356
Qy	901	ACCTCTCGGTGATGATCCAGGTCGCTTCCTTCAGCTGACGGGACGCGCGTCTTCCGT	960
Db	13357	AGCTTTCTGTGATCTCTCAGGTGGCTTCCTTTAAACTGCGCGGACACGATATTTTCGC	13416
Qy	961	ATGGCGCCGATCCATCACCATTTCGAACCTGAAAGGCTGGCCGACCCGCGCGTGTATCGTG	1020
Db	13417	ATGGCACCGATTTCATCACCACCTATGAACCTGAAAGGCTGGCCGGAACCGCGTCAATTGTG	13476
Qy	1021	CGCTTCTGATCATCACCGTGATCTCTGTGCTGATCGGCTCGCACCTTGAAGCTGGT	1080
Db	13477	CGTTCTCGAATTATTCGCTGATGCTGGTCTCTGATTGGTCTGGCAACGCTGAAGGTACGT	13536
Qy	1081	TGA 1083	
Db	13537	TAA 13539	
RESULT 14			
AX191720			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
FEATURES			
source			
ORIGIN			
Query Match			
Best Local Similarity			
Matches			
Qy			
Db			
Qy			
Db			
Qy			
Db			
Qy			
Db			
Qy			
Db			
Qy			
Db			
Qy			
Db			
Qy			
Db			
Qy			
Db			
Qy			
Db			
Qy			
Db			
Qy			
Db			
Qy			
Db			
Qy			
Db			
Qy			
Db			
Qy			
Db			
Qy			
Db			
Qy			
Db			
Qy			
Db			
Qy			
Db			
Qy			
Db			
Qy			
Db			
Qy			
Db			
Qy			
Db			
Qy			
Db			
Qy			
Db			
Qy			
Db			
Qy			
Db			
Qy			
Db			

QY	361	ATCGAGAAGAACTCCCGTGGCCCTGCGAGCCGCTGGAAGTACTTCTGGCAGTGGGTTC	420
Db	12892	GTGCGTAAAGACACCAAGAGGTTGATCGCTGTTGGAAGTATTCTTCGATGTCGGTCATT	12951
QY	421	GGCATGGCGCCGCGCTGTTCTCTACATGACCTGCGAACCCTGATCGAGACCCCTG	480
Db	12952	GCCTGGGTGTCGCTTCGCCCTTGTTACCTTGCAGCAAGACACGCCCAACGACGAGCTG	13011
QY	481	ATCGTCCCGATGTCGAAGACGCTGAGATCCAGTTGGGCATCTTCTTCGTGCTCTGACC	540
Db	13012	GTGGTCCCATCTTTAAAGATGTGATGCGCAGCTGGGGCTGTTCTACATCTCTGCTGCT	13071
QY	541	TACTTGGTATCGCTCGCTCGACATGGGTAACTCACTACCGAGTCTCGACGCCCTG	600
Db	13072	TACTTGGTATCGCTCGCTCGACATGGGTAACTCACTACCGAGTCTCGACGCCCTG	13131
QY	601	GCATCATGCGCAGCTTAATGTTGCGCGCGCTCGGCATCTTCTGCTACCTGTCTGGGC	660
Db	13132	GCAATTATGCGGACCGTATTGTCGCGGTGTTGTCGCTGGTGGCGACCGGC	13191
QY	661	AACGTGAGTTCCGAGTACCTGCTGATTCCTCAAGTACCGGCGCCGCGGAGCTGATC	720
Db	13192	AATGAACTTTGCCAGCTACTTGCATATACCGTATCTCGACACCGCGGGAACCTGTT	13251
QY	721	GTGTTCTGCGCGCGCTGGTGGCGCGCGCTCGGCTTCTCTGCTTCAACACCTATCG	780
Db	13252	ATTGCTGTACCGGATAGTCGGGCGAGACTGGGCTTCTGTGGTTTACACCTATCG	13311
QY	781	GCGAGTCTTCAATGGCGAGCTGCGCGCGCTGGCGCTGGCGCCCGCTGGGACCACTC	840
Db	13312	GCGAGTCTTCAATGGCGAGCTGCGCGCGCTGGCGCTGGCGCCCGCTGGGACCACTC	13371
QY	841	GCGGTGATCGTGGCGCAGGAGTCTGCTGCTTCACTATGCTGGTGGGTTCGTCATGGA	900
Db	13372	GCGTACTGCTAGTCAAGAACTCTGCTGCTGATGATGCGGCGCGCTGCTGCTGGA	13431
QY	901	ACCTCTCGTGTATGATCCAGTCTGCTTCTTCAAGTCACTGACGCGCGCTCTTCCTG	960
Db	13432	ACGTTCTCTGATCTCTGCGAGTCTGCTTCTTAACTGCGCGCAACAGTATTTTCG	13491
QY	961	ATGCGCGGATCATCACTTCACTTCACTGAAAGGCTGCGGACCGCGCGCTGATCGT	1020
Db	13492	ATGCGCGGATCATCACTTCACTTCACTGAAAGGCTGCGGACCGCGCGCTGATCGT	13551
QY	1021	CGCTTCTGATCATCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1080
Db	13552	CGTTCTGATTTATTCGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	13611
QY	1081	TGA 1083	
Db	13612	TAA 13614	
RESULT 15			
EC2MIN	standard; genomic DNA; PRO; 28277 BP.		
XX	AC	X55034; M10429;	
XX	SV	X55034.1	
XX	DT	21-FEB-1991 (Rel. 27, Created)	
XX	DT	05-JUL-1999 (Rel. 60, Last updated, Version 31)	
XX	DE	E. coli 2 minute region	
XX	KW	ddl gene; envA gene; ftsA gene; ftsQ gene; ftsW gene; ftsZ gene; ilvH gene;	
XX	KW	ilvI gene; leuA gene; leuO gene; mraY gene; murC gene; murD gene;	
XX	KW	murE gene; murF gene; murG gene; mutT gene; orfA; orfB; orfC; orfX;	
XX	KW	pdpB gene; secA gene; shl gene.	
OS	XX	Escherichia coli	

OC	Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;	
OC	Enterobacteriaceae; Escherichia.	
XX	[1]	
RP	1-28277	
RA	Ayala J.A.;	
RT	Submitted (08-JAN-1991) to the EMBL/GenBank/DBJ databases.	
RL	Ayala J.A.; Instituto de Biologia Molecular, Centro de Biologia Molecular,	
RL	Universidad Autonoma, Canto-Blanco 28049, Madrid, Spain.	
XX	[2]	
RA	Ayala J.A.;	
RT	"Regulation of transcription at the 2-minute region of the genetic map of	
RT	Escherichia coli.";	
XX	Unpublished.	
XX	[3]	
RP	1-306	
RA	MEDLINE; 82078077.	
RT	PUBMED; 6171647.	
RA	Wessler S.R.; Calvo J.M.;	
RT	"Control of leu operon expression in Escherichia coli by a transcription	
RT	attenuation mechanism";	
RL	J. Mol. Biol. 149(4):579-597(1981).	
XX	[4]	
RP	268-1130	
RA	MEDLINE; 86223773.	
RT	PUBMED; 3519576.	
RA	Haughn G.W., Wessler S.R., Gemmill R.M., Calvo J.M.;	
RT	"High A + T content conserved in DNA sequences upstream of leuABCD in	
RT	Escherichia coli and Salmonella typhimurium";	
RL	J. Bacteriol. 166(3):1113-1117(1986).	
XX	[5]	
RP	843-1812	
RA	MEDLINE; 88320486.	
RT	PUBMED; 3413113.	
RA	Henikoff S., Haughn G.W., Calvo J.M., Wallace J.C.;	
RT	"A large family of bacterial activator proteins";	
RL	Proc. Natl. Acad. Sci. U.S.A. 85(18):6602-6606(1988).	
XX	[6]	
RP	1799-2187	
RA	MEDLINE; 85234358.	
RT	PUBMED; 3891724.	
RA	Haughn G.W., Squires C.H., DeFelice M., Largo C.T., Calvo J.M.;	
RT	"Unusual organization of the ilvIH promoter of Escherichia coli";	
RL	J. Bacteriol. 163(1):186-198(1985).	
XX	[7]	
RP	2101-4431	
RA	MEDLINE; 83272971.	
RT	PUBMED; 6308579.	
RA	Squires C.H., DeFelice M., Devereux J., Calvo J.M.;	
RT	"Molecular structure of ilvIH and its evolutionary relationship to ilvG in	
RL	Escherichia coli";	
XX	Nucleic Acids Res. 11(15):5299-5313(1983).	
XX	[8]	
RP	4274-6093	
RA	MEDLINE; 90330585.	
RT	PUBMED; 2198273.	
RA	Leclerc G., Noel G., Drapeau G.;	
RT	"Molecular cloning, nucleotide sequence and expression of shl, a new gene	
RL	in the 2-minute region of the genetic map of Escherichia coli";	
XX	J. Bacteriol. 172(8):4696-4700(1990).	
XX	[9]	
RP	6088-7587	
RA	MEDLINE; 90251464.	
RT	Gomez M.J., Fluoret B., Van Heijenoort J., Ayala J.A.;	

RT "Nucleotide sequence of the regulatory region of pbpB gene of *Escherichia coli*";  
 RL Nucleic Acids Res. 18:2813-2813(1990).  
 XX [10]  
 RN 7316-10074  
 RP MEDLINE; 83296957.  
 RX PUBMED; 6350821.  
 RA Nakamura M., Maruyama I.N., Soma M., Kato J.I., Suzuki H., Hirota Y.;  
 RT "On the process of cellular division in *Escherichia coli*: nucleotide  
 RL Mol. Genet. 191(1):1-9(1983).  
 XX [11]  
 RN MEDLINE; 90124047.  
 RP PUBMED; 2692800.  
 RX Tao J.S., Ishiguro E.B.;  
 RA "Nucleotide sequence of the *murE* gene of *Escherichia coli*";  
 RL Can. J. Microbiol. 35(11):1051-1054(1989).  
 XX [12]  
 RN MEDLINE; 90328986.  
 RP PUBMED; 2198024.  
 RX Michaud C., Parquet C., Flouret B., Blanot D., Van Heijenoort J.;  
 RA "Revised interpretation of the sequence containing the *murE* gene encoding  
 RT the UDP-N-acetylmuramyl-tripeptide synthetase of *Escherichia coli*";  
 RL Biochem. J. 269(1):277-278(1990).  
 XX [13]  
 RN 11142-12634  
 RP MEDLINE; 89345095.  
 RX Parquet C., Flouret B., Menguin-Lecreux D., Van Heijenoort J.;  
 RA "Nucleotide sequence of the *murF* gene encoding the UDP-MurNac-  
 RT synthetase of *Escherichia coli*";  
 RL Nucleic Acids Res. 17:5379-5379(1989).  
 XX [14]  
 RN 12423-15030  
 RP MEDLINE; 90192099.  
 RX Ikeda M., Wachi M., Ishino F., Matsubashi M.;  
 RA "Nucleotide sequence involving *murD* and an open reading frame *orf-Y* spacing  
 RT *murF* and *ftsW* in *Escherichia coli*";  
 RL Nucleic Acids Res. 18:1058-1058(1990).  
 XX [15]  
 RN 13392-15020  
 RP Menguin-Lecreux D., Van Heijenoort J.;  
 RA "Nucleotide sequence of the *murD* gene encoding the UDP-MurNac-  
 RT synthetase of *Escherichia coli*";  
 RL Nucleic Acids Res. 18:183-183(1990).  
 XX [16]  
 RN 14743-16239  
 RP MEDLINE; 90036736.  
 RX Ikeda M., Sato T., Wachi M., Jung H.K., Ishino F., Kobayashi Y.,  
 RA "Structural similarity among *Escherichia coli* *PtsW* and *RodA* proteins and  
 RT *Bacillus subtilis* *SpoVE* protein, which function in cell division, cell  
 RL elongation, and spore formation, respectively";  
 RL J. Bacteriol. 171(11):6375-6378(1989).  
 XX [17]  
 RN 16094-17806  
 RP Menguin-Lecreux D., Texier L., Van Heijenoort J.;  
 RA "Nucleotide sequence of the cell-envelope *murG* gene of *Escherichia coli*";  
 RL Nucleic Acids Res. 18:2810-2810(1990).  
 XX [18]  
 RN 16094-18886  
 RP MEDLINE; 90326550.  
 RX Ikeda M., Wachi M., Jung H.K., Ishino F., Matsubashi M.;  
 RT "Nucleotide sequence involving *murG* and *murC* in the *mra* gene cluster region

RT of *Escherichia coli*";  
 RL Nucleic Acids Res. 18:4014-4014(1990).  
 XX [19]  
 RN 19464-21952  
 RP MEDLINE; 85054557.  
 RX PUBMED; 6094474.  
 RA Robinson A.C., Kenan D.J., Hatfull G.F., Sullivan N.F., Spiegelberg R.,  
 RT "DNA sequence and transcriptional organization of essential cell division  
 RL genes *ftsQ* and *ftsA* of *Escherichia coli*: evidence for overlapping  
 RL transcriptional units";  
 RL J. Bacteriol. 160(2):546-555(1984).  
 XX [20]  
 RN 21464-23333  
 RP MEDLINE; 86083166.  
 RX PUBMED; 3000876.  
 RA Yi Q.M., Lutkenhaus J.;  
 RT "The nucleotide sequence of the essential cell-division gene *ftsZ* of  
 RL *Escherichia coli*";  
 RL Gene 36(3):241-247(1985).  
 XX [21]  
 RN 22964-25011  
 RP MEDLINE; 88058745.  
 RX PUBMED; 2824434.  
 RA Beall B., Lutkenhaus J.;  
 RT "Sequence analysis, transcriptional organization, and insertional  
 RL mutagenesis of the *envA* gene of *Escherichia coli*";  
 RL J. Bacteriol. 169(12):5408-5415(1987).  
 XX [22]  
 RN 23989-27799  
 RP MEDLINE; 88298644.  
 RX PUBMED; 2841285.  
 RA Schmidt M., Rollo E., Grodberg I., Oliver D.;  
 Query Match 46.4%; Score 502.2; DB 15; Length 28277;  
 Best Local Similarity 66.5%; Pred. No. 2.3e-56;  
 Matches 720; Conservative 0; Mismatches 363; Indels 0; Gaps 0;  
 QY 1 ATGCTCTGCTGCTGGCCGAATACCTGCAACAGTTCTACAGGGCTTCGGCGCTTTCCAG 60  
 DB 12532 ATGTTAGTTGGCTGGCCGAACATTTGGTCAATATTATTCGGCTTTAACGCTTTTCC 12591  
 QY 61 TACCTGACCTCGCGGGCAATTCAGCGTGCTCAGCGGCTGTCGCTGCTGCTGCTGCTG 120  
 DB 12592 TATCTGACGTTTCGGGCCATCGCTGAGCTGCTGAGCCGCTCTTCATCTCATTTGGATG 12651  
 QY 121 GGGCCCTGGATGATCGTACCTTGCGAGATCCCGGAGATCCCGGCGCTGGCAACGAC 180  
 DB 12652 GGCCCGGTATGATGCTCATTTGCAAAACTTTCTTGGTTCAGGTGGTGGTAAACGAC 12711  
 QY 181 GTTCCGACGTCGACCTGTGGAAGAAGGCACCCCGACCATGGCGGCGCCCTGATCCTT 240  
 DB 12712 GGTCTGTAATCACATTCAGCAAGCGGTACCGGACCATGGCGGGAATTATGATCTGT 12771  
 QY 241 ACGCCATAGCCATCAGACGCTGCTGGGGGATCTTTCCAAACGCTACGCTGTTGGTA 300  
 DB 12772 ACGCGAATTTGATCTCCGCTACTGCTGTTGGGTTTACCCTGTCATCCGTACCTGTTGC 12831  
 QY 301 GTGCTGCTGTTACCTGCTGTTTCGGTGCCATCGGCTGGGTAGACGACTACCGCAAGGTG 360  
 DB 12832 GTGTTGTTGTTGCTGTGTAGTTTACGTTATTGGCTTTGTTGATGATTATCGCAAAGTG 12891  
 QY 361 ATCGAAGAACTCCCGTGGCTGCGGACCCGCTGGAGTACTTCTGCGAGTGGGTGTC 420  
 DB 12892 GTGCGTAAAGACACCAAGGGTTGATCGCTCTGTTGAAGTATTCTTGGATGTCGGTCA 12951  
 QY 421 GGCATCGGCGCGCGCTGTTTCCTCTACATGATCGCGAAACCCCGATCGAGACCCCTG 480  
 DB 12952 GCGCTGGGTGCTGCGCTTCGCCCTGTACCTTCCGCGCAAGACACGCCCGCAACGACGCTG 13011

481	ATCTGTCGCGATCGCTGAGAGCGTCAGATCCAGTTGGGATCTTCTTCGTGGTCCCTGACC	540
13012	GTGGTCCCAATCTTTAAAGATGTGANGCGCAGCTGGGGCTGTTCATCATTTCTCTGGCT	13071
541	TACTTCGTTCATCGTCGGCTCGAGCAATGCGGTGAACCTCACCGAGGTCTCGACGGGCTG	600
13072	TACTTCGTTCATTTGGGTACTGGCAACGCGTAAACCTTGACCGATGGTCTCGACGGGCTG	13131
601	GGGATCATCGCACGGTAATGTTGCCGGCGCTGGGCATCTTCGTGCTACCTCTCGGGC	660
13132	GCAATTATGCCGACGTAATTTGTCCCGGTGGTTTTGCGTGTGGCGTGGCGACCGGC	13191
661	AAGCTGAATTCCGCCAGTACCTGTGATTCGCAACGTACCGGGCGCGCGGAGCTGATC	720
13192	AATATGAACATTTGCCAGCTACTTGCATATATCCGTATCTCGACACGCCGGGGAACCTGGTT	13251
721	GTGTCTCGCGCGCGCTGGTCGGCGCCGGCTCGGCTTCCTCTGCTTCAACACACTATCCG	780
13252	ATTGCTGTGACCGCATAGTCGGGGCAGGACTGGGCTTCCTGTGGTTAAACCTATCCG	13311
781	GGCGAGTCTTCATGGGAGAGTGGCGCGCTGGCGCTGGCGCGCGCGCTGGGACCATC	840
13312	GGCGAGTCTTTATGGGCGATGATGGTTCTGGTGGCGTTAGTGTGGCTTAGGCAATTATC	13371
841	GGCGTGATCGTGGCGCAGAGATCGTGTCTGTTCACTATGGGTGGGGTGTTCGTATCGAA	900
13372	GCGCTACTGCTACGTACAGAAATTCCTGTGGTGAATTATGGGGGCGTGTTCGTGGTAGAA	13431
901	ACCCCTCGGTGATGATCCAGGTGCTTCCTTCAGCTGACGGACGCGCGCTCTTCGGT	960
13432	AGCTTTCTGTTCATCTGCAGGTGGCTTCCTTTAAACTCGCGGACAACTGATTTTCCGC	13491
961	ATGGGCGGATCCATCAACATTTGCAACTGAAGGCTGGCGGACCCGCGCGTGAATCGTG	1020
13492	ATGGCACCGAATTCATCACCACTATGAACTGAAAGGCTGGCGGGAACCGCGCGTCAATTGTG	13551
1021	CGTCTCTGATCATCACCGTGTATCTGTGTCTGATCGGCGCTCGCACCTTGAAGCTCGGT	1080
13552	CGTTTCGTGATTATTCGCTGATGTGATTCTGATGGTCTGGCAACGCTGGAAGGTAGGT	13611
1081	TGA	1083
13612	TAA	13614

Search completed: May 9, 2004, 14:51:52  
Job time : 4400 secs